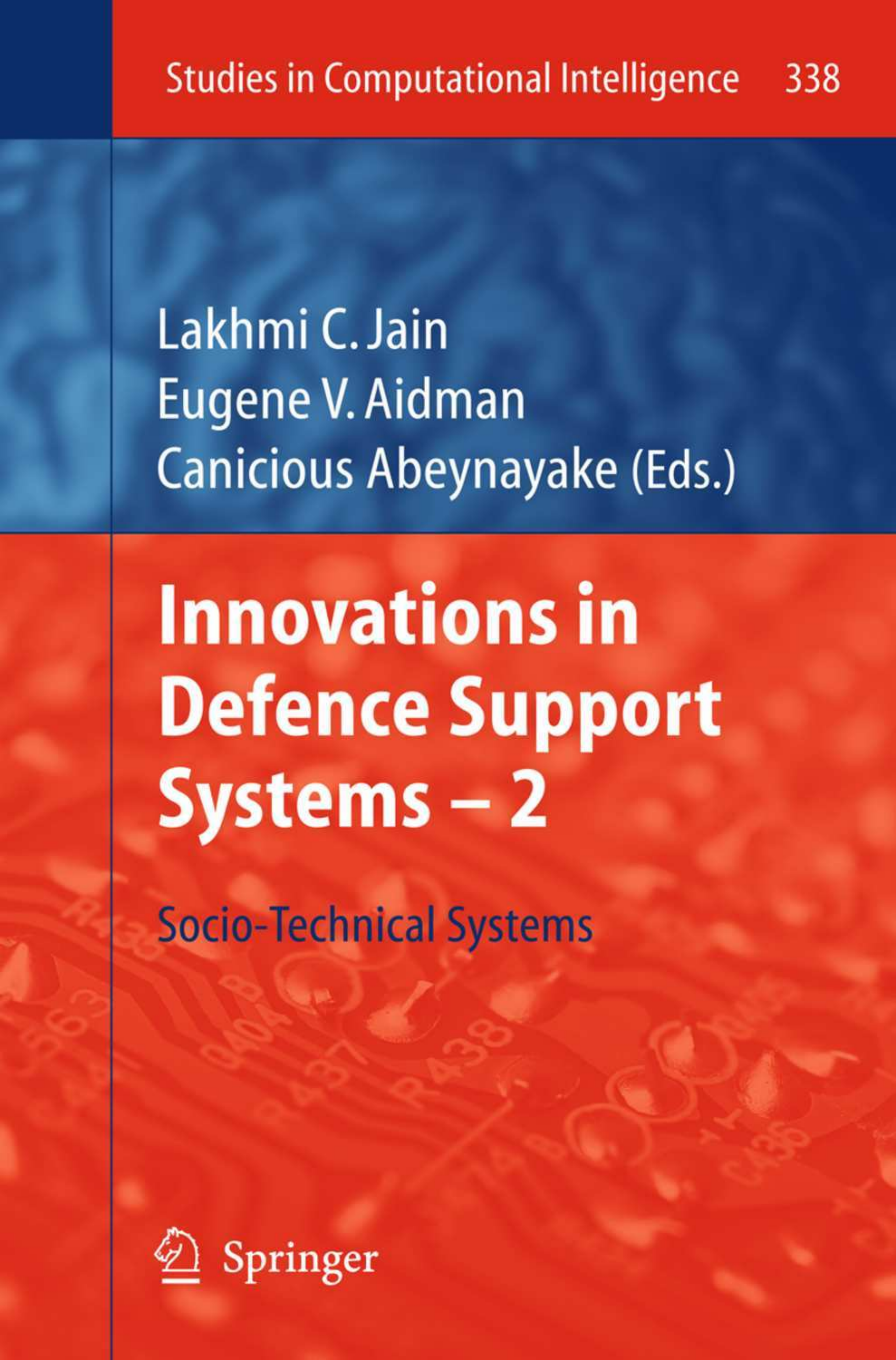
Studies in Computational Intelligence 338



Lakhmi C. Jain
Eugene V. Aidman
Canicious Abeynayake (Eds.)

Innovations in Defence Support Systems – 2

Socio-Technical Systems

Springer

Lakhmi C. Jain, Eugene V. Aidman, and Canicious Abeynayake (Eds.)

Innovations in Defence Support Systems – 2

Studies in Computational Intelligence, Volume 338

Editor-in-Chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 312. Patricia Melin, Janusz Kacprzyk, and
Witold Pedrycz (Eds.)
Soft Computing for Recognition based on Biometrics, 2010
ISBN 978-3-642-15110-1

Vol. 313. Imre J. Rudas, János Fodor, and
Janusz Kacprzyk (Eds.)
Computational Intelligence in Engineering, 2010
ISBN 978-3-642-15219-1

Vol. 314. Lorenzo Magnani, Walter Carnielli, and
Claudio Pizzi (Eds.)
Model-Based Reasoning in Science and Technology, 2010
ISBN 978-3-642-15222-1

Vol. 315. Mohammad Essaaidi, Michele Malgeri, and
Costin Badica (Eds.)
Intelligent Distributed Computing IV, 2010
ISBN 978-3-642-15210-8

Vol. 316. Philipp Wolfrum
Information Routing, Correspondence Finding, and Object Recognition in the Brain, 2010
ISBN 978-3-642-15253-5

Vol. 317. Roger Lee (Ed.)
Computer and Information Science 2010
ISBN 978-3-642-15404-1

Vol. 318. Oscar Castillo, Janusz Kacprzyk,
and Witold Pedrycz (Eds.)
Soft Computing for Intelligent Control and Mobile Robotics, 2010
ISBN 978-3-642-15533-8

Vol. 319. Takayuki Ito, Minjie Zhang, Valentin Robu,
Shaheen Fatima, Tokuro Matsuo,
and Hirofumi Yamaki (Eds.)
Innovations in Agent-Based Complex Automated Negotiations, 2010
ISBN 978-3-642-15611-3

Vol. 320. xxx

Vol. 321. Dimitri Plemenos and Georgios Miaoulis (Eds.)
Intelligent Computer Graphics 2010
ISBN 978-3-642-15689-2

Vol. 322. Bruno Baruque and Emilio Corchado (Eds.)
Fusion Methods for Unsupervised Learning Ensembles, 2010
ISBN 978-3-642-16204-6

Vol. 323. Yingxu Wang, Du Zhang, and Witold Kinsner (Eds.)
Advances in Cognitive Informatics, 2010
ISBN 978-3-642-16082-0

Vol. 324. Alessandro Soro, Vargiu Eloisa, Giuliano Armano,
and Gavino Paddeu (Eds.)
Information Retrieval and Mining in Distributed Environments, 2010
ISBN 978-3-642-16088-2

Vol. 325. Quan Bai and Naoki Fukuta (Eds.)
Advances in Practical Multi-Agent Systems, 2010
ISBN 978-3-642-16097-4

Vol. 326. Sheryl Brahnam and Lakhmi C. Jain (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare 5, 2010
ISBN 978-3-642-16094-3

Vol. 327. Slawomir Wiak and
Ewa Napieralska-Juszczak (Eds.)
Computational Methods for the Innovative Design of sElectrical Devices, 2010
ISBN 978-3-642-16224-4

Vol. 328. Raoul Huys and Viktor K. Jirsa (Eds.)
Nonlinear Dynamics in Human Behavior, 2010
ISBN 978-3-642-16261-9

Vol. 329. Santi Caballé, Fatos Xhafa, and Ajith Abraham (Eds.)
Intelligent Networking, Collaborative Systems and Applications, 2010
ISBN 978-3-642-16792-8

Vol. 330. Steffen Rendle
Context-Aware Ranking with Factorization Models, 2010
ISBN 978-3-642-16897-0

Vol. 331. Athena Vakali and Lakhmi C. Jain (Eds.)
New Directions in Web Data Management 1, 2011
ISBN 978-3-642-17550-3

Vol. 332. Jianguo Zhang, Ling Shao, Lei Zhang, and
Graeme A. Jones (Eds.)
Intelligent Video Event Analysis and Understanding, 2011
ISBN 978-3-642-17553-4

Vol. 333. Fedja Hadzic, Henry Tan, and Tharam S. Dillon
Mining of Data with Complex Structures, 2011
ISBN 978-3-642-17556-5

Vol. 334-336. xxx

Vol. 337. Sheryl Brahnam and Lakhmi C. Jain (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare 6, 2011
ISBN 978-3-642-17823-8

Vol. 338. Lakhmi C. Jain, Eugene V. Aidman, and
Canicious Abeynayake (Eds.)
Innovations in Defence Support Systems – 2, 2011
ISBN 978-3-642-17763-7

Lakhmi C. Jain, Eugene V. Aidman, and
Canicious Abeynayake (Eds.)

Innovations in Defence Support Systems – 2

Socio-Technical Systems

Springer

Prof. Lakhmi C. Jain
University of South Australia
School of Electrical and
Information Engineering
SA 5095 Adelaide
South Australia
Australia

Dr. Eugene V. Aidman
School of Psychology,
University of Adelaide
South Australia, SA 5005
Principal Staff Officer Science,
HQ Forces Command
Paddington NSW 2021
Australia

Dr. Canicious Abeynayake
Weapons Systems Division
Defence Science and Technology
Organisation
B190 EOP
P.O. Box 1500
Edinburgh, S.A. 5111
Australia

ISBN 978-3-642-17763-7 e-ISBN 978-3-642-17764-4

DOI 10.1007/978-3-642-17764-4

Studies in Computational Intelligence ISSN 1860-949X

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The rapidly advancing warfighting technology has lead to a recognition by the military of the critical need for a paradigm shift to place the human in the centre of military systems development. Both conventional system design and the artificial intelligence paradigms have converged on the principles of user-centred design, resulting in a range of break-through solution types - from sophisticated operator interfaces for conventional systems, based on augmented cognition and assistive technologies, to autonomous and semi-autonomous platforms aspiring at replacing – or partially replacing – the human effort. The emerging theme in most – if not all – of these inventions is that the object of design is not a technical system that gets adjusted to suit the operator, but a holistic socio-technical system that has human input at its core, in a specification-defining capacity.

We wish to publish the state-of-art research on defence related systems in our future volumes of the sub-series as the previous volumes on defence systems have proved useful not only to the research community but to a wider community of researchers in the field.

The focus of the current volume is on the design and optimization of such socio-technical systems and their performance in defence contexts. Conceptual and methodological considerations for the development of such systems and criteria likely to be useful in their evaluation are discussed, along with their conceptual underpinnings in total system performance analysis.

The book has assembled contributions from leading academics and defence scientists and represents a “state of the art” in their respective fields.

The book is directed to researchers, practitioners and advanced graduate students in systems engineering and human factors.

Thanks are due to the authors and reviewers for their contribution. The editorial support by the Springer-Verlag is acknowledged.

Lakhmi C. Jain
Adelaide, Australia

Eugene V. Aidman
Sydney, Australia

Canicious Abeynayake
Adelaide, Australia

Contents

Chapter 1

Advances in Defence Support Systems 1
Lakhmi C. Jain, Eugene Aidman
1 Introduction 1
2 Chapters Included in the Book 2
3 Conclusion 3
References and Further Reading 4

Chapter 2

Multi-Robot Task Allocation for Performing Cooperative Foraging Tasks in an Initially Unknown Environment 5
Prithviraj Dasgupta
1 Introduction 5
2 Existing Techniques for MRTA 6
3 Multi-Robot Task Allocation for Foraging Tasks 7
4 Dynamic Multi-Robot Task Allocation Algorithm 8
    4.1 Local Heuristics for Task Selection 11
5 Experimental Results 12
    5.1 Experimental Setup 12
    5.2 Experiments to Analyze the MRTA Algorithm and Task Selection Heuristics 14
6 Conclusions and Future Work 17
References 18

Chapter 3

Formal-Language-Oriented Foundation of Dynamics of Human Crowds Using Topos-Theory 21
Vladimir G. Ivancevic, Darryn J. Reid
1 Introduction 21
2 A Brief on Category Theory 30
    2.1 Maps 30
    2.2 Topological Spaces 32

- 2.3 Commutative Diagrams 36
- 2.4 Categories 38
- 2.5 Functors 41
- 2.6 Natural Transformations 44
- 2.7 Limits and Colimits 47
- 2.8 Adjunction 47
- 2.9 Groups and Related Algebraic Structures 49
- 2.10 Snake Lemma and Tensor Products 53
- 2.11 A Brief on Categorical Logic 55

3 A Brief on Topos Theory 57
- 3.1 Topoi: Mathematical Universes with Intuitionistic Logic 57
- 3.2 The Topos of Varying Sets 60
- 3.3 Sets through Time 62
- 3.4 Presheaves on a Poset 63
- 3.5 Presheaves on a General Category 67
- 3.6 Classical Realism vs. Quantum Instrumentalism 71

4 Propositional Languages and Crowd Dynamics 72
- 4.1 Three Interpretations of Propositions 72
- 4.2 The Propositional Language $\mathcal{PL}(S)$ 73
  - 4.2.1 Intuitionistic Logic and the Definition of $\mathcal{PL}(S)$ 73
  - 4.2.2 Representations of $\mathcal{PL}(S)$ 75
  - 4.2.3 The Representation of $\mathcal{PL}(S)$ in Crowd Mechanics 77
  - 4.2.4 The Failure to Represent $\mathcal{PL}(S)$ in Crowd Behavior 78
- 4.3 Instrumentalism, Tableau and Resolution 79
  - 4.3.1 Instrumentalism 79
  - 4.3.2 Analytic Tableau 81
  - 4.3.3 Resolution 82

5 A Higher-Order, Typed Language for Crowd Behavior 82
- 5.1 The Basics of the Language $\mathcal{TL}(S)$ 82
- 5.2 Representing $\mathcal{TL}(S)$ in a Topos 86
- 5.3 Crowd Mechanics in the Local Language $\mathcal{TL}(S)$ 88
- 5.4 Adapting the Language $\mathcal{TL}(S)$ to Other Types of Crowd System 89

6 The Category of Complex Systems 90
- 6.1 The Category Sys 91
  - 6.1.1 The Arrows and Translations for the Disjoint Sum $S_1 \sqcup S_2$ 91
  - 6.1.2 The Arrows and Translations for the Composite System $S_1 \otimes S_2$ 93
  - 6.1.3 The Concept of ‘Isomorphic’ Systems 94
  - 6.1.4 An Axiomatic Formulation of the Category Sys 95
- 6.2 Representations of Sys in Topoi 98

6.3 Crowd Mechanics in Sys 100
6.3.1 Details of the Translation Representation 102
7 General Crowd Dynamics in a General Topos 104
7.1 The Pull-Back Operations 104
7.1.1 The Pull-Back of Physical Quantities 104
7.1.2 The Pull-Back of Propositions 106
7.2 The Topos Rules for General Crowd Dynamics 107
References 110

Chapter 4

Methodology for the Evaluation of an International Airport Automated Border Control Processing System 115
Veneta MacLeod, Brett McLindin
1 Introduction 115
2 Defining the Problem Space and Scoping the Analysis 118
2.1 Manual Primary Line Processing 120
2.2 Automated Border Control System Processing 120
2.3 Defining the Components within the System Boundary 121
2.3.1 Manual Border Control System Boundary 122
2.3.2 Automated Border Control System Boundary 122
3 Determining the Quantitative and Qualitative Measures of Performance 125
3.1 Technical Performance 126
3.2 Matching Performance 126
3.3 Process Timings 127
3.4 Observations 128
3.5 Traveller's Perceptions 129
4 Collecting and Processing the Data 131
4.1 Ground Truth 131
4.2 Matching Performance 131
4.3 Process Timings and Observations 132
4.4 Traveller's Perceptions 136
5 Analysing and Interpreting the Data Output 137
5.1 Ground Truth 137
5.2 Matching Performance 138
5.3 Process Timings 139
5.4 Observations 142
5.5 Traveller's Perceptions 143
6 Reporting the Results 143
7 Conclusion 144
References 144

Chapter 5

The Role of the Human Operator in Image-Based Airport Security Technologies 147

Ian Graves, Marcus Butavicius, Veneta MacLeod, Rebecca Heyer, Kathryn Parsons, Natalie Kuester, Agata McCormac, Philip Jacques, Raymond Johnson

1 Introduction 147
2 Image Based Security Technologies in the Airport Environment 148
  2.1 The Airport Environment 148
  2.2 Image Based Security Technologies 150
    2.2.1 Face Recognition Systems 151
    2.2.2 Baggage Screening Systems 151
    2.2.3 Passenger Screening Systems 152
3 The Role of the Human Operator within the System 154
  3.1 Automation of Security Tasks 154
    3.1.1 Reasons for Automation 155
    3.1.2 Limitations of Fully Automated Systems 156
  3.2 The Impact of the Human Operator on Overall Performance 156
    3.2.1 Consequences of Non-optimal Decision Thresholds 156
    3.2.2 Limitations to the Impact of the Human Operator 157
4 Factors That Impact Human Operator Performance 158
  4.1 Usability 158
  4.2 Training 158
  4.3 Workload 159
    4.3.1 Number of Alarms 159
    4.3.2 Multi-tasking 160
    4.3.3 Individual versus Team Dynamics 161
  4.4 Fatigue and Stress 161
  4.5 Human Operator Individual Differences 162
  4.6 Biases 163
    4.6.1 Ethnicity, Gender and Age Biases in Face Recognition 164
    4.6.2 Cognitive Bias 164
    4.6.3 Reliance on Automation 165
    4.6.4 Consequence of Actions 166
    4.6.5 Speed of Processing 166
5 Recommended Methodologies for Evaluation of Systems 167
  5.1 Types of Evaluations 167
    5.1.1 Technical Assessments 167
    5.1.2 Scenario Tests 167
    5.1.3 Operational Trials 168
  5.2 Primary Measures of Performance 168
    5.2.1 Detection Accuracy 169
    5.2.2 Discrimination and Bias 169
    5.2.3 Subjective Confidence 170

5.2.4 Processing Speed ... 170
5.2.5 Cognitive, Perceptual and Other Abilities ... 171
5.2.6 Usability Assessment ... 171
6 Challenges in Evaluating Human Operator Performance ... 172
6.1 Methodology Design ... 173
6.2 Sample Size Considerations ... 174
6.3 Context of the Results ... 175
7 Conclusions ... 175
References ... 176

Chapter 6

Assessment of the ThruVision T4000 Passive Terahertz Camera: A Human Factors Case Study ... 183
M.A. Butavicius, K.M. Parsons, A. McCormac, R. Foster, A. Whittenbury, V. MacLeod
1 Background ... 183
1.1 Previous DSTO Study ... 184
1.2 Previous Research Outside DSTO ... 185
1.3 Theoretical Background ... 188
1.3.1 Group Interaction ... 188
1.3.2 Feedback and Learning ... 189
1.3.3 Rational Experiential Inventory ... 189
1.3.4 Visual Processing ... 190
1.4 Summary of Research Aims ... 191
2 Methodology ... 192
2.1 Participants ... 192
2.2 Materials ... 192
2.2.1 Demographic Questionnaire ... 192
2.2.2 Personality and Cognitive Abilities ... 192
2.2.3 THz Camera Footage ... 194
2.3 Procedure ... 194
3 Results ... 196
3.1 Signal Detection Theory Measures ... 196
3.2 Observations on B’’ and Conversion to |B’’| ... 197
3.3 REI Scores ... 197
3.4 The Influence of One versus Two Operator Conditions ... 197
3.5 The Influence of Feedback ... 198
3.6 The Influence of Personality on Performance ... 200
3.7 The Influence of Visual Abilities on Performance ... 200
4 Discussion and Conclusions ... 201
4.1 One versus Two Operator Conditions ... 201
4.2 The Influence of Feedback ... 201
4.3 The Influence of Personality Factors ... 202
4.4 The Influence of Perceptual Factors ... 202

4.5 Limitations ... 202
4.6 Future Directions ... 203
References ... 204

Chapter 7

Biorhythmic Analysis to Prevent Aviation Accidents ... 207

R.K. Saket, Wg. Cdr. S.P. Kaushik, Col. Gurmit Singh

1 Introduction ... 208
2 Typical Biorhythm Cycles: An Overview ... 210
2.1 Interpretation of Biorhythm Cycles ... 212
2.2 Accident Prone Zone / Critical Days ... 212
2.3 Combined Biorhythm Cycles ... 213
3 The Causes of Aircraft Accidents ... 213
3.1 Direct Causes of Aircraft Accidents ... 214
3.1.1 Technical Defects in Aircraft ... 214
3.1.2 Environment Factors ... 214
3.1.3 Human Factors ... 215
3.2 Indirect Causes of Aircraft Accidents ... 215
3.2.1 Cumulative Stress ... 215
3.2.2 Effects of Sudden Stress ... 216
4 Behavioral Analysis and the Impact of Cycles of Biorhythm on Human Performance: Bio Analysis ... 216
4.1 Behavior Based Biorhythm Management ... 219
4.2 Causes of an Accidents and Safety Measures ... 219
5 Analysis of the Causes of Human Error Factor and the Development of a Model for Predicting Human Error ... 221
5.1 Man Machine Interaction When Considering Flying Mistakes ... 222
5.2 An Error Taxonomic Approach ... 223
5.3 Factors Which Affect Decision Making ... 223
5.4 Factors Affecting Sensitivity ... 224
5.5 Factors Affecting Criterion ... 224
5.6 Rationale for Research on Visual Inspection Training ... 225
5.6.1 Visual Lobe Training ... 226
5.6.2 Feedback Training ... 226
5.6.3 Feed Forward Training ... 227
5.6.4 Attribute Training ... 227
5.6.5 Schema Training ... 228
5.6.6 Interaction of SRK Behavior ... 228
6 Statistical Technique for Biorhythmic Analysis ... 228
7 Reliability Evaluation of Biorhythmic Aviation Accidents ... 230
7.1 Gaussian Distribution Approach ... 230
7.2 Safety Factor Concept and Peak Demand Considerations ... 232
7.3 Performance Evaluation Based on Peak Demand Using Simpson 1/3rd Rule ... 233

8 Results and Discussion 234
9 Conclusion 237
References 238

Chapter 8

Reliability Evaluation of Defence Support Systems 241
R.K. Saket
1 Introduction 241
2 Probability Theory: An Overview 244
  2.1 Conditional Probability 244
  2.2 Independent Events 246
  2.3 Baye's Theorem 246
  2.4 Unsolved Numerical Problems 249
3 Hazard Model of Defence Support Systems 250
4 Life Distributions of Defence Support System: An Overview 250
  4.1 The Binomial Distribution 251
  4.2 The Exponential Distribution 251
  4.3 Poisson Distribution 251
  4.4 Geometric Distribution 254
  4.5 The Weibull Distribution 255
  4.6 The Normal Distribution 256
  4.7 The Gamma Distribution 259
5 Basic Concepts of System Reliability 259
6 Mean Time to Failure (MTTF) of Defence Components 261
7 Reliability of Defense Support System Structures 262
  7.1 Series Systems 262
  7.2 Parallel Systems 264
  7.3 Series-Parallel (k-out-of-n) Systems 265
8 Failure Distribution Functions and Reliability of IDSS 266
  8.1 Reliability of Exponential Distribution Systems 266
  8.2 Reliability of Weibull Distribution Systems 270
  8.3 Reliability of Normal Distribution Systems 273
  8.4 Reliability and MTBF Evaluation Using Exponential Model 275
    8.4.1 Series Connected Components 275
    8.4.2 Parallel Connected Components 275
  8.5 Miscellaneous Solved Numerical Problems 276
9 Conclusion 284
References 285

Author Index 287

Chapter 1
Advances in Defence Support Systems

Lakhmi C. Jain[1] and Eugene Aidman[2]

[1] School of Electrical and Information Engineering,
University of South Australia,
Mawson Lakes Campus,
South Australia SA 5095,
Australia

[2] School of Psychology, University of Adelaide,
South Australia, SA 5005
Principal Staff Officer Science, HQ Forces Command
Paddington NSW 2021
Australia

Abstract. This chapter previews the main themes and contributions to this volume, focusing on the design and optimization of socio-technical systems and their performance in defence contexts. Conceptual and methodological considerations for the development of such systems and criteria likely to be useful in their evaluation are discussed, along with their conceptual underpinnings in total system performance analysis.

1 Introduction

The tremendous advances in the conventional system design techniques as well as in the artificial intelligence paradigms have generated continuously growing interest in the design of adaptive and robust systems for defence.

The rapidly advancing warfighting technology has lead to a recognition by the military of the critical need for a paradigm shift to place the human in the centre of military systems development:

> *It is important to always remember in dealing with advanced technology, especially when it comes to warfighting, the Soldier imperative. It is all about the soldiers. We will never replace people on the battlefield. They are the people who make the technology work, and we can never forget that.* (Lynch, 2001, p.2).

The current emphasis on collective, team and distributed systems performance adds to the complexity of military systems. This complexity is further magnified by the free-play, adversarial nature of the interactions they have to support, which leads to a need for new methods capable of capturing the essence and mechanisms of this complexity. Chapters in this volume will discuss some of

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 1–4.
springerlink.com

these new methods, and the overall philosophy of systems-oriented human factors research that underpins human-centric systems design, which will be illustrated with some examples of focused practical applications.

The other important theme this volume is attempting to address is systems integration. It is often the most challenging component of complex and multifaceted projects (Unewisse, Straughair, Pratt, Kirby & Kempt, 2003). The difficulties encountered in the systems integration of a number of major defence projects, such as the Collins Class Submarine, Sea Sprite helicopters, and Wedgetail have highlighted the risks associated with large scale systems integration. These examples of challenging systems integration have been undertaken in the relatively benign environment, of a single platform environment, in the context of integrated projects.

Defence support systems are particularly challenging for systems integration. They include multiple independent nodes dispersed across a large and/or complex environment, using austere and possibly intermittent communications, a requirement to regroup the component elements during the course of an operation, continuously operating in an environment, which is inherently hostile. Individual warfighting capabilities are typically produced by a large number of separate projects, each focused on optimising their capabilities, with integration into the wider force a secondary consideration. The new Defence systems will not only need to integrate with each other but with a variety of legacy systems. As a consequence, the systems integration of the future forces is likely to prove far more challenging than the systems integration of individual platforms.

The integration of the individual platforms and legacy capabilities into a network-enabled force will be a major challenge in system integration across the whole range of inputs, with a particularly strong emphasis on the development of the human dimension of warfighting. The consdiderations here will range from concepts, doctrine, tactics and training, to personnel selection and development, leadership and cultural change. This book presents new approaches, tools and techniques aimed at addressing the issues if human-centric system integration.

2 Chapters Included in the Book

This book includes eight chapters. Chapter 1 provides an introduction to the defence support systems.

Chapter 2 examines the performance of robotic teams in cooperative foraging tasks. It presents a novel approach to multi-robot task allocation for optimizing the performance of robotic teams in these cooperative tasks in an initially unknown environment.. The tasks require target location to be discovered and for multiple robots to share the task's execution to successfully complete it. A novel, emergent algorithm is developed using a set of heuristics to select the order of the tasks. The Webots simulator is used to validate the proposed heuristics and to analyze their relative performance. The proposed techniques provide a simple, scalable and efficient mechanism for allocating dynamically arriving tasks across multiple robots, which offers promise in the development of autonomous and semi-autonomous defence platforms and their concepts of operation.

Chapter 3 proposes a novel mathematical formulation for a general dynamics of human crowds. The approach can be applied both to relatively simplistic crowd mechanics such as the physical motion of individual agents, aggregates and a crowd as a whole, as well as to complex cognitive characteristics and behaviors of a crowd, which incorporate motivation and cognition into the crowd mechanics. A quantum-probability based approach to modelling such crowd dynamics is presented.

Chapter 4 is on a methodology for the evaluation of an international airport automated border control processing system. The methodology represents a tested protocol that can be used for the evaluation of other automated biometric systems within different operational environments.

Chapter 5 is on the role of human operator in image-based airport security technologies. The authors present a number of human factors issues which will have an impact on human operator performance in the operational environment, as well as highlighting the variables which must be considered when evaluating the performance of these technologies in different scenarios.

Chapter 6 is on the assessment of the ThruVision Passive T4000 Passive Terahertz Camera. The purpose of this research is to assess a number of the human factors issues surrounding the use of this device in an airport scenario. The authors present the theoretical background, methodology (including psychological pre-tests) as well as aspects of the results from a controlled psychological experiment.

Chapter 7 examines the utility of biological rhythms monitoring and analysis in the prevention of aviation accidents. The authors present a new technique based on probabilistic approach to biorhythmic analysis to prevent aviation accidents.

Chapter 8 reviews the well-established concept of reliability and introduces a new discipline of reliability evaluation in engineering and technological design for defence support systems. The Chapter covers a range of topics including the basic concepts of probability theory, reliability evaluations for defence support system structures, a Hazard model for failure analysis and various probability distributions involved in reliability evaluation, including both solved and unsolved numerical examples.

3 Conclusion

This chapter has presented a sample of the summary of research work undertaken by various research groups in the area of defence support systems. It is obvious that intelligent paradigms play a significant role in modern defence support systems.

References and Further Reading

[1] Finn, A., Jain, L.C. (eds.): Innovations in Defence Support Systems-1. Springer, Heidelberg (2010)

[2] Finn, A., Scheding, S.: Developments and Challenges for Autonomous Unmanned Vehicles. Springer, Heidelberg (2010)

[3] Jain, L.C., Ichalkaranje, N., Tonfoni, G. (eds.): Advances in Intelligent Systems for Defence. World Scientific Publishing Company, Singapore (2002)

[4] Balbo, F., Pinson, S.: A Transportation Decision Support System in Agent-based Environment. Intelligent Decision Technologies 1(3), 97–115 (2007)

[5] Lambert, D., Scholz, J.: Ubiquitous Command and Control. Intelligent Decision Technologies 1(3), 157–173 (2007)

[6] Silverman, B., et al.: An Embeddable Testbed for Insurgent and Terrorist Agent Theories. Intelligent Decision Technologies 2(4), 195–203 (2008)

[7] Wickler, G., Potter, S.: Information Gathering: From Sensor Data to Decision Support in Three Simple Steps. Intelligent Decision Technologies 3(1), 3–17 (2009)

[8] Gorodetsky, V., et al.: Agent-based Distributed Decision Making in Dynamical Operational Environments. Intelligent Decision Technologies 3(1), 35–57 (2009)

[9] Marchese, M., et al.: An Interaction-Centric Approach to Support Peer Coordi-nation in Distributed Emergency Response Management. Intelligent Decision Technologies 3(1), 19–34 (2009)

[10] Cumming, M.L., Bruni, S.: Human-Automated Planner Collaboration in Complex Resource Allocation Decision Support Systems. Intelligent Decision Technologies 4(2), 101–114 (2010)

[11] Taghezout, N., Zarate, P.: Supporting a Multicriterion Decision Making and Multi-agent Negotiation in Manufacturing Systems. Intelligent Decision Technologies 3(3), 139–155 (2009)

[12] Ren, X., et al.: An Agent-based System for Distributed Fault Diagnosis. International Journal of Knowledge-based and Intelligent Engineering Systems 10(5), 319–335 (2006)

[13] Yang, M., Shen, Q.: Reinforcing Fuzzy Rule-based Diagnosis of Turbo-machines with Case-based Reasoning. International Journal of Knowledge-based and Intelligent Engineering Systems 12(2), 173–181 (2008)

[14] Unewisse, M., Straughair, P., Pratt, J., Kirby, B., Kempt, N.: Land 5000: The integration of land force combat capabilities. In: Proceedings of the Land Warfare Conference, Adelaide, Australia, October 28-30, pp. 295–309 (2003)

[15] Lynch, R.: Lessons Learned: Commanding a Digital Brigade Combat Team.: Institute For Defence Analyses, Alexandria, VA, Paper P-3616 (2003)

Chapter 2

Multi-Robot Task Allocation for Performing Cooperative Foraging Tasks in an Initially Unknown Environment

Prithviraj Dasgupta

Computer Science Department,
University of Nebraska, Omaha, NE 68182, USA
pdasgupta@mail.unomaha.edu

Abstract. We consider the problem of multi-robot task allocation (MRTA) by a set of robots that are deployed in an initially unknown environment to perform foraging tasks. The location of each task has to be discovered by each robot through searching in the environment. Each task also requires multiple robots to share the task's execution to complete the task. We discuss an emergent, distributed MRTA algorithm and describe a set of heuristics that can be used by the MRTA algorithm to select the order of the tasks so that the tasks are performed in an efficient manner. The heuristics are evaluated using simulated robots on the Webots simulator to analyze their relative performance.

1 Introduction

Over the past few years, multi-robot task allocation (MRTA) has emerged as an important direction in robotics research. The MRTA problem can be broadly defined as the following - given a set of tasks, a set of robots that can perform the tasks, and an objective function that measures the performance efficiency of different combinations of robots from the robot set in performing the tasks, find a suitable matching or allocation between the set of tasks and the set of robots which optimizes the value of the objective function. The MRTA problem is encountered in different application domains of multi-robot systems including unmanned search and rescue, cooperative transportation, autonomous exploration and mapping, distributed monitoring and surveillance, etc. In this chapter, we consider the problem of solving a dynamic MRTA problem where the temporal and spatial distribution of tasks is not known *a priori* by the robots. Specifically, we describe a cooperative foraging problem where robots have to search for tasks in the environment. When a task gets discovered by a robot, multiple other robots have to share the execution the task to complete it. The MRTA problem specifies the coordination rules used by the robots in such a scenario to perform the tasks in an efficient manner. The MRTA problem in a cooperative foraging scenario with shared task execution by multiple robots has been shown to be NP-hard in our previous work. We describe a distributed MRTA algorithm using an aggregation

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 5–20.
springerlink.com

strategy between robots. We then illustrate four different heuristic functions which have increasing levels of sophistication, and simultaneously increasing complexity. These heuristics can be used by the MRTA algorithm running on each robot to select a task to execute, so that the tasks are performed in an efficient manner. We have empirically compared the efficiency of performing tasks while using these heuristics for MRTA in a distributed target recognition scenario where robots have to cooperatively perform visual identification of tasks that are spatially and temporally distributed in an environment.

The rest of this chapter is structured as follows. In the next section, we discuss the related work in the field of multi-robot task allocation. Section 3 describes the MRTA problem within the framework of foraging with dynamically arriving tasks. Experimental results with multiple robots for a distributed target identification application within the Webots robot simulation platform are given in Section 4, and finally, we conclude.

2 Existing Techniques for MRTA

The problem of multi-robot task allocation (MRTA) has been investigated using different techniques [12, 27, 28, 37, 38] over the last decade. One of the earliest systems to use market-based algorithms for MRTA was the M+ system described by Botelho and Alami [5]. In this system, information about all tasks is assumed to be available to all robots. A contract-net based protocol [33] followed by a consensus protocol to guarantee mutual exclusion is used by each robot to select the lowest cost task it can perform. Gerkey [12] also provides an important and widely accepted classification taxonomy for MRTA problems. The problems are classified along three dimensions: (a) single task (ST) versus multi-task(MT), related to the parallel task performing capabilities of robots, (b) single robot(SR) versus multi-robot(MR), related to the number of robots required to perform a task, and, (c) instantaneous assignment(IA) versus time extended assignment(TA), related to the planning performed by robots to allocate tasks. Mataric *et al.* [22] compared the performance of robots teams within an emergency scenario of putting out alarms in a 2-d environment and showed that the least time is required to complete all tasks (put out all alarms) when the robots are allowed to coordinate their plans with each other as well as dynamically change their plans. The *traderbots* approach by Dias [11] is another significant milestone that uses auction-based MRTA. Traderbots uses multi-round, single-item auctions for dynamic task allocation across multiple robots. A recent extension of the traderbots approach [16] uses heterogeneous robot teams for performing tightly coordinated tasks by combining the traderbots approach for task allocation with the Skill, Tactics, Play (STP) approach for coordinated teamwork. Lagoudakis *et al.* [18, 19] address MRTA as an exploration problem of matching a set of robots to a set of targets and provide theoretical bounds on the performance of their allocation algorithm called PRIM-ALLOCATION. Sariel and Balch [29, 30] have improved the PRIM-ALLOCATION algorithm using an auction-based solution for the Multi-TSP problem. In [39], Zheng *et al.* provide an improved algorithm for multi-round auctions to attempt to move closer in performance to combinatorial auctions without the computational complexity using two mechanisms - a larger lookahead and a reinforcement learning inspired technique called rollout. More recently, Zlot [40] has described several auction-based algorithms for multi-robot task allocation. The work focuses mainly on task tree decomposition of loosely coupled tasks that can be specified using AND-OR relationships. Different auction based algorithms, both centralized and decentralized, for finding efficient allocations of decomposed tasks across robots are described using different objective functions such as minimizing cost and makespan. Other MRTA approaches are given in [4, 6, 15, 17, 21, 25, 26]. In [14, 20], the authors have used single-item auction algorithms for MRTA on unmanned aerial vehicles (UAVs) performing search and reconnaissance tasks. Recently, the problem of MRTA has been combined with a negotiation framework to improve the allocations through possible task reallocations [13, 32, 36]. Tang and Parker [34] describe an auction-based coalition formation algorithm for MRTA with robots equipped with different types of sensors used for a box pushing experiment. Yet another recent enhancement of the market-based MRTA algorithms has been the introduction of learning techniques such as reinforcement learning [31] and vector regression learning [16] to enable a robot learn from its own schedule history and predict the expected future reward from a task so that it can bid more efficiently on the task.

Many of the papers discussed above consider scenarios where each task can be completed by a single robot and robots do not need to coordinate with each other to perform a single task. In contrast, in this chapter we consider a task allocation and execution model where coordinated actions from multiple robots are required to complete each task, and robots can commit to multiple tasks simultaneously. We also consider model the effect of partial actions by robots on a task. Our problem can therefore be classified as the more complex SR-MT-TA setting described in [12] where a robot can perform only one task at a time(SR), multiple robots are required to complete a task(MT) and robots store the tasks assigned to them as a task list within them and adjust the order of tasks in the task list to plan their paths efficiently(TA).

3 Multi-Robot Task Allocation for Foraging Tasks

We consider a foraging problem domain where wheeled mobile robots have to search and discover objects of interest in a two-dimensional environment individually, but have to collaborate with each other to perform actions on the objects of interest. A detailed description of the environment is given in [10]. Here, we provide an overview of the salient features of the system and its environment. The set of actions required to be performed on an object of interest by a robot is called a task. Because each object of interest in our environment is associated with the set of tasks performed on it by robots, we have used the term 'task' to refer to objects of interest in the rest of the chapter. We have purposefully kept the notion of a task abstract at this point in the chapter as the set of actions to be performed on an object is largely application specific. In Section 5, we illustrate a very simple search-and-execute scenario where a task performed by a robot corresponds to simple color identification of the object it discovers. Following are some of the important features of our system:

1. Tasks are distributed randomly within the environment. The spatial distribution of tasks is not known *a priori* by the robots. The robots perform a distributed search in the environment to discover the tasks. A robot needs to move to the vicinity of a task to be able to discover it by perceiving/observing it on its sensors. In this chapter, we consider stationary tasks only.
2. A single robot is only capable of discovering and partially performing tasks but lacks the computational resources required to complete a task on its own. A task can be completed only if multiple robots collaborate to share their computational resources towards executing the task. We consider loosely coupled tasks and different robots collaborating with each other to perform the same task can execute those actions asynchronously and independently of each other.
3. To enlist the cooperation of other robots required to complete a task, a robot that discovers a task communicates the task's information to other robots.
4. On completing its portion of execution of a task a robot communicates the progress of the task after its execution (fraction of task still incomplete) to other robots within its communication range. Those robots then consider their own commitments and selectively choose to visit the task to perform it. After completing its portion of a task, a robot either continues to perform any other tasks it is already committed to, or reverts to individually searching the environment for tasks if it does not have any other committed tasks.

The multi-robot foraging scenario described above involves different phases including the deployment strategy of the robots within the environment, coverage and exploration strategy used by the robots, the MRTA algorithm used by the robots, and finally, the algorithm used by the robots to execute tasks. In the rest of this chapter, we have focused only on the MRTA algorithm that is used to allocate a task discovered by a robot between the other robots, so that the task can be completed in an efficient manner. The strategies for the remaining phases of the foraging problem are described in [10].

In [24], we have shown that the MRTA problem for a cooperative foraging domain with dynamically arriving tasks is NP-complete, by reducing it to a dynamic traveling salesman problem. In the next section, we describe the distributed MRTA algorithm for the cooperative foraging scenario, which uses different heuristics to calculate an approximate solution to this problem.

4 Dynamic Multi-Robot Task Allocation Algorithm

In a cooperative foraging scenario, the tasks to be performed can arrive dynamically and the locations of the tasks are not known *a priori* by the robots. Therefore, the robots have to first search for tasks in the environment[1]. When a task gets

---

[1] Our focus in this chapter is on the MRTA algorithm between robots and we have used a simple Braitenberg motion algorithm on each robot to enable it to cover the environment while searching for tasks. In Braitenberg motion, a robot continues to move in a straight line until it encounters an obstacle perceived through its sensors. The robot then turns at an appropriate angle to avoid colliding with the obstacle and continues its straight line motion. More advanced algorithms for distributed area coverage in a cooperative foraging scenario are described in [7, 8].

discovered by a robot, the robot performs or executes its portion of the task. Subsequently, other robots have to share the execution of the discovered task in an asynchronous manner, and ensure that the task gets completed efficiently. The MRTA algorithm gives the rules of coordination between the robots required to achieve this shared execution of tasks. To realize the MRTA algorithm in a distributed manner we have used different aggregation strategies that are implemented using local heuristics on each robot's controller. The distributed MRTA algorithm also maintains and updates information about the tasks in the environment that have been discovered and performed by robots.

The task information communicated between robots after discovering a new task or after performing a task that was discovered by another robot, forms the backbone of the MRTA algorithm. The task information consists of the location of the task in 2D coordinates as perceived by the robot that discovers it, and a real-number representing the portion of the task completed by the cumulative action of the robots on the task. We have used a real number similar to a virtual pheromone value used in many emergent systems [3, 35] to represent the portion of a task that is completed. Pheromone values are allowed to evaporate or decay with time to represent the volatility of operations performed the environment. A task is considered complete when the amount of virtual pheromone associated with it reaches a certain threshold value.

The MRTA algorithm uses three data structures to record information about the tasks it is aware of. The first task-related data structure is the *allocatedTaskList* that contains information about tasks that have been discovered and communicated by other robots and have to be completed. The elements in the allocatedTaskList are ordered using the value of one of the heuristic functions described in Section 4.1. The *visitedTaskList* is an unordered list that contains information about the tasks that have been already visited by the robot. Finally, the *completedTaskList* contains information about tasks that have been completed by robots and should be ignored if they are re-discovered. Figure 1 shows the data structures and their update procedure used by the MRTA algorithm. The *allocatedTaskList* and *completedTaskList* are updated when a robot receives task-related communication from other robots. If the communicated task information is related to a new task that is not on any of the task lists of the robot, the MRTA algorithm adds the task to its *allocatedTaskList* and recalculates the heuristic function value for each task in the *allocatedTaskList* using one of the heuristic functions described in Section 4.1. If the communicated task information is related to the update (and not task completion) of a task in the allocatedTaskList, the information related to the task is updated in the *allocatedTaskList*. If the task update information is related to a task in the *visitedTaskList* or *completedTaskList*, it is ignored. On the other hand, if the communicated task information is related to a task that has just been completed, the MRTA algorithm adds the task information to its *completedTaskList*. The *allocatedTaskList* and *visitedTaskList* are also checked for the task, and if it is present in either of these lists, the task is removed from the list. The *visitedTaskList* is updated when the robot visits the location of a task, either while searching for tasks in the environment and discovering the task first hand, or by visiting the task after being allocated to perform it using the MRTA algorithm.

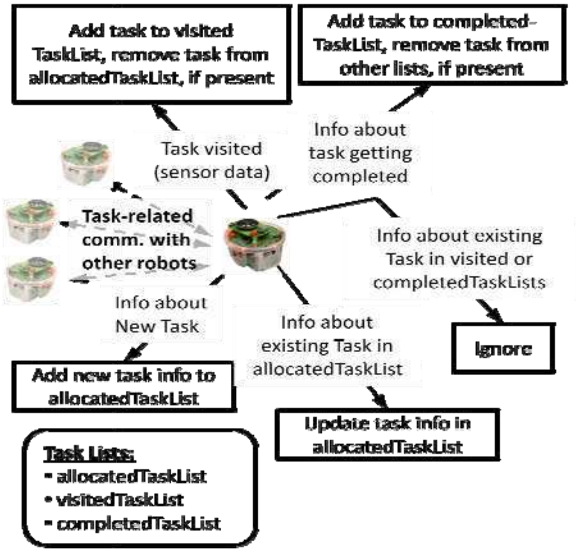


Fig. 1. Different task lists and task list management functions done by the distributed MRTA algorithm

The state diagram of the robot's controller for performing tasks within the cooperative foraging scenario while using the MRTA algorithm is shown in Figure 2. To select a task from its *allocatedTaskList*, the MRTA algorithm

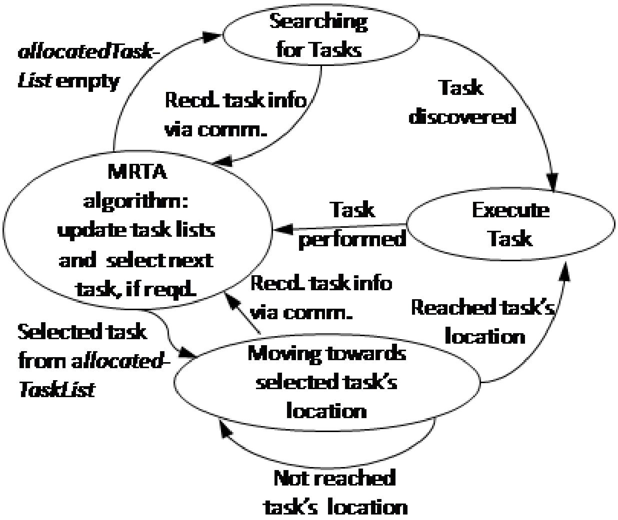


Fig. 2. State diagram of a controller of a robot performing cooperative foraging using the distributed MRTA algorithm

determines the task with the highest value according to the particular heuristic function being used. The robot then starts to move towards the selected task's location. While moving towards the task's location, if the robot receives information that the task it is moving towards is completed, it stops moving towards the task and selects the next task from its allocatedTaskList using the heuristic function. Otherwise, the robot continues its motion until it reaches the location of the task and performs it. The MRTA algorithm then checks the allocatedTaskList to select the next task. If the allocatedTaskList is empty, the robot reverts to searching for tasks in the environment.

4.1 Local Heuristics for Task Selection

To address the complexity of the distributed MRTA problem in the cooperative foraging setting, we have used different multi-robot aggregation strategies. Each strategy is implemented using a heuristic function that calculates the suitability or priority of a task in the robot's allocatedTaskList. The different heuristic functions that have been used for this purpose are described below:

- ***Closest Task First (CT):*** Each robot selects a task from its task list that is closest to it. Distances are normalized over the sum of distances to all tasks in a robot's task list to enable comparison between task distances. The value of this heuristic is very simple to calculate. However, it can lead all robots to prefer tasks that are close to them and consequently give a lower preference to tasks further away from them. Consequently, tasks that are further away from most of the robots can remain incomplete for long times and more than the required number of robots can get allocated to tasks that are closer to most of the robots. Overall, the closest task first heuristic is not very efficient in spatially distributing the task load across the robots.
- ***Most Starved Task First (MST):*** To address the drawbacks of the closest task first heuristic and to balance the task load across the environment, it would make sense for robots to give a higher priority to tasks that have the least number of robots in their vicinity and are likely to be requiring more robots to complete them. The most starved task first heuristic does this by enabling each robot to select a task from its task list that has the least number of robots in its vicinity.
- ***Most Starved, Most Complete Task First (MSMCT):*** A potential drawback of the most starved task first heuristic is that robots are attracted towards a recently discovered task which is likely to have few robots near it. This can result in almost complete tasks being left incomplete because robots prefer more starved and possibly less complete tasks over a task that is almost nearing completion but has more robots (which possibly already visited the task) near it. To address this problem, the most starved, most complete task first heuristic extends the most starved task first heuristic by considering the number of robots still required to complete a task. While selecting a task, a robot using this heuristic considers a product of the number of robots in the vicinity of the task as well as the progress of the

task from the pheromone value associated with the task. Tasks that are nearing completion are given higher preference.

- ***Most Proximal Task First (MPT):*** In the previous two heuristics, the MRTA algorithm only considers the effect of other robots on tasks in its *allocatedTaskList,* but it does not consider the robot's own relative position to the other robots. This can result in robots that are further away from the task allocating the task and moving towards it, only to be informed en-route that other robots have completed the task before it. To alleviate this problem, the most proximal task first heuristic first determines how many other robots are closer to the task being considered than the robot itself. It then selects the task that has the least number of robots closer to the task than itself and also is the nearest towards being completed.

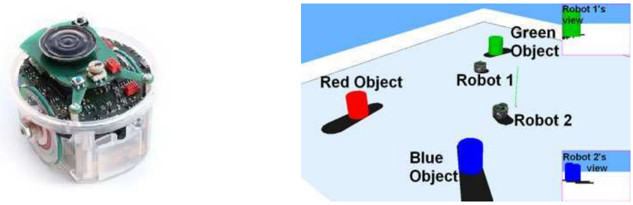
5 Experimental Results

We have measured the efficiency of our heuristic-based MRTA algorithm in terms of the time required to complete the tasks in the environment, as well as in terms of the task-load faced by each robot and the number of robots performing each task. Taken together, the time required by the robots to perform tasks and the allocation between robots and tasks reflect the amount of energy that is expended by the robots to perform the tasks in the environment.

5.1 Experimental Setup

We have verified our distributed MRTA algorithm empirically using different environment settings and different number of robots within a commercially available robot simulator called Webots [23]. Webots provides a powerful simulation platform for accurately modeling complex environments and multiple robots including their physics. Each robot is simulated as an e-puck robot (shown in Figure 3(a)) that is a two-wheeled robot with a dsPIC-based processor capable of 14 MIPS, 144 KB Flash and 8 KB RAM, and measuring 7 cm in diameter. It has the following sensors: (1) Camera: a color VGA camera with a maximal resolution of 640 X 480. (2) Eight infra-red distance sensors measuring ambient light and proximity of obstacles in a range of 4 cm. (3) A Bluetooth-enabled transmitter and receiver for sending and receiving messages between robots. To provide localization capability to the robots, we have added a GPS node on each e-puck robot within the Webots simulation. For all our experiments, the environment is considered as a 3.5 X 3.5 m^2 square region. A snapshot of our test environment within Webots with two e-puck robots and three colored targets is shown in Figure 3(b).

We have used a cooperative, distributed target identification application as our test scenario. The scenario consists of a set of stationary targets. Each target is a cylindrical object with a unique color. Targets are distributed randomly in the environment and the objective is to identify all the targets while reducing the time required and the energy expended by each robot to identify targets. The location of the targets is not known *a priori* by the robots and must be discovered in real time. Each robot uses a Braitenberg controller to navigate in the environment, which allows them to avoid obstacles and avoid each other using their distance sensors. In this scenario, a *task* corresponds to the actions performed by a robot on a target. For example, a single robot's task is to execute an image (color) identification algorithm when it observes a colored object of interest on its camera. Based on the color of the object of interest, the robot classifies it as a possible target. The robot also associates $\Pi_g = 20$ units of virtual pheromone with the location corresponding to the object of interest. The virtual pheromone decays exponentially according to the equation Π_g X $\exp^{k(t-t_c)}$, where t is the time at which the pheromone was last updated and t_c is the current time. A suitable value for the constant k that balances the volatility of targets with the velocity and distribution of robots was experimentally determined to be $k=1.5$ X 10^{-4}. To confirm a recently discovered object as a target, a certain number of additional robots need to cooperate by visiting the object's location, observing the object and successfully executing their respective image identification algorithms on the object, and concurring with the initial robot's classification of the object as a target. To achieve this multi-robot coordination, when a robot discovers an object of interest, it sends a broadcast message to other robots that are within its communication range. The broadcast message includes the location and pheromone information of the object. The robots that receive this message then use the MRTA algorithm described in this chapter to decide whether to visit the location of the object and perform their respective actions on it, thereby increasing the amount of virtual pheromone associated with the object. An object is considered to be confirmed as a target (in other words, a task is considered to be complete) when the cumulative virtual pheromone deposited at it by different robots reaches a threshold value Π_{Thr}. For our experiments we have used $\Pi_{Thr} > 60$. This means that with $\Pi_g = 20$, a task is completed when at least n_r = ceiling (Π_g / Π_{Thr}) = 4 robots visit the target. However, the virtual pheromone associated with a task's location is allowed to decay over time. Therefore, if a significant amount of time elapses between the task's discovery by a robot and the task's shared execution by other robots, more than 4 robots can be required to complete the task. In some extreme cases of inefficient task allocation, tasks can also never get completed within a finite time if other robots are not able to visit the task's location to perform it before the pheromone deposited at the task's location decays to a nominal value. All the experimental results reported in this section were averaged over 20 simulation runs.

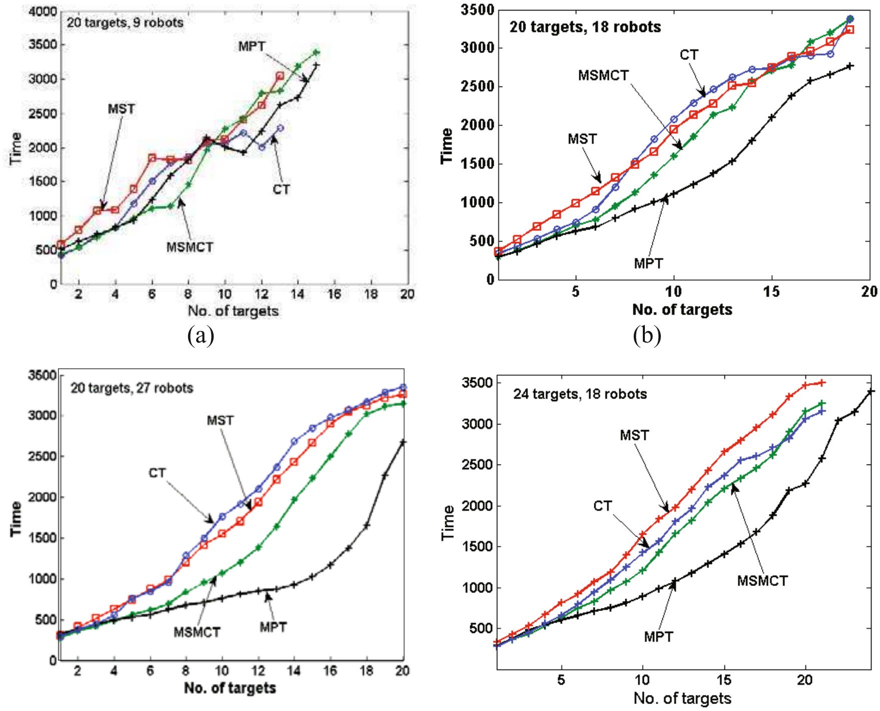


(a) (b)

Fig. 3. (a) Photograph of an e-puck robot (Courtesy: www.gctronic.com). (b) A scenario inside the Webots simulator showing two e-puck robots and three objects, each with a different color.

5.2 Experiments to Analyze the MRTA Algorithm and Task Selection Heuristics

Our first experiment analyzes the effect of using the different heuristics-based aggregation strategies described in Section 4.1 with different number of robots. The environment for these sets of simulations consists of 20 targets with 9, 18 or 27 robots. We have used the time required to confirm all the targets in the environment as the metric for comparing the strategies. For each of these configurations, we observe that the most proximal task first heuristic outperforms the other strategies. In Figure 4(a), with 20 targets and 9 robots, the robots perform comparably while using the most proximal task first (MPT) heuristic, albeit slightly better than when the robots use any of the other heuristics. However, with only 9 robots in the environment and each task requiring at least n_r=4 robots to get completed, some of the tasks are perennially starved, and, consequently, the robots are able to complete 15 out of the 20 tasks in the environment. This situation is ameliorated when the number of robots in the environment is increased to 18 and 27 respectively, while keeping the number of targets fixed at 20. Figure 4(b) shows that with 18 robots in the scenario, the robots are able to perform all the tasks in 23.6% lower time while using the MPT heuristic than while using any of the other heuristics. A comparable performance advantage of 23.1% is obtained with the MPT heuristic when the number of robots is increased to 27, as shown in Figure 4(c). Finally, Figure 4(d) shows that when the number of targets is increased to 24 with 18 robots in the environment, while using the MPT heuristic the robots are still outperform the performance obtained using the other heuristics by about 24%. With the MPT heuristic, the robots are able to distribute the tasks more efficiently and can consequently finish all the tasks in the environment. On the other hand, with the other heuristics, the robots are able to complete only 20 out of the 24 tasks. The superior performance of the MPT heuristic can be attributed to the fact that in the MPT heuristic, a robot considers its relative ‘fitness’ towards the progress and completion a task in comparison to that of other robots that have a higher possibility of performing the task, because of the latter’s closer proximity to the task. Robots can therefore reduce inefficient allocations where a robot selects a task only to discover that other robots that had selected the task before it were able to complete the task, after the robot started moving towards the task but before the robot was able to reach it. The other heuristics do not consider the relative ‘fitness’ of a robot with respect to other robots. Even between the remaining heuristics, we observe that as amount of information used by the heuristic to make the task selection decision increases, the heuristics perform correspondingly better.



(a) (b)

(c) (d)

Fig. 4. Time required to confirm all targets (complete all tasks) for the different heuristics-based aggregation strategies in the MRTA algorithm with different numbers of robots. (a) 20 targets, 9 robots, (b) 20 targets, 18 robots, (c) 20 targets, 27 robots, (d) 24 targets, 18 robots.

For our next set of experiments, we analyzed the task load of the robots for different numbers of robots and tasks (targets). For this set of experiments, a scenario with 18 robots and 20 targets was considered as the baseline scenario and both the number of robots and targets were varied to observe their effect on the robot's task load. Figure 5(a) shows that as the number of robots increases while keeping the number of targets fixed at 20, each robot is required to visit fewer targets. This happens because as the number of robots increases, the task load gets shared between more robots and consequently, each robot's load decreases. On the other hand, if the number of robots is kept fixed the load of each robot increases with the number of tasks, as shown in Figure 5(b). Between the heuristics, the MPT heuristic gives the highest task load for each robot. To understand this, we first recall from the results in Figure 4(b) that when the robots use the MPT heuristic they are able to complete tasks in less time as compared to when they use other heuristics. Therefore, with the MPR heuristic, the robots are more available to do the tasks and consequently can take on more tasks to perform.

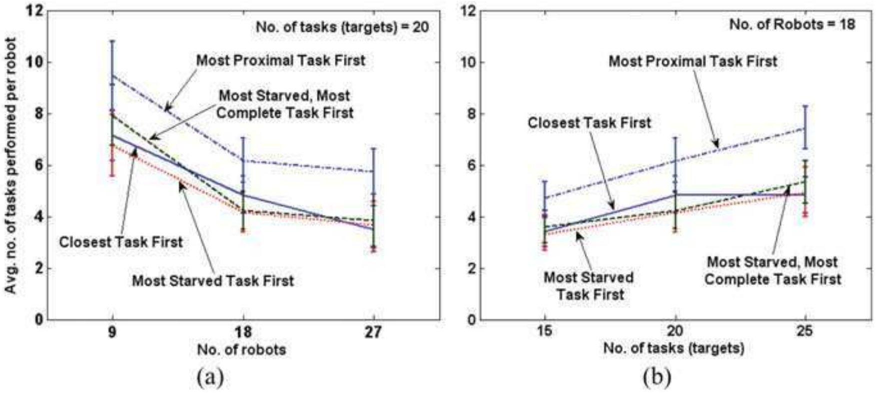
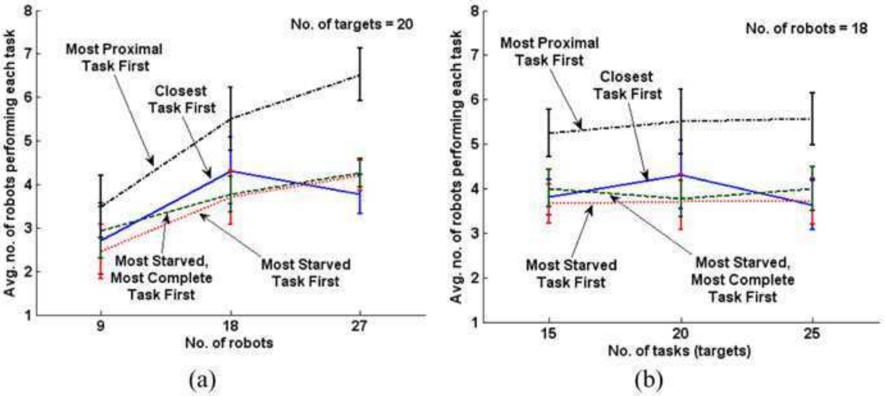


Fig. 5. Average number of tasks performed (targets visited) by each robot for the different aggregation heuristics used. (a) For different numbers of robots, (b) For different numbers of tasks (targets).

For our final set of experiments, we analyze the number of robots that visit a target to understand how efficiently tasks are performed by robots. As mentioned in Section 5.1, for the pheromone values we have used in our experiments, a minimum of $n_r = 4$ robots are required to complete a task. This value of $n_r = 4$ then represents the optimum number of robots visiting a target. Values of n_r higher than 4 represent more than the desired number of robots getting allocated to a task, while n_r values below 4 represent tasks remaining incomplete. Once again we vary the number of robots and number of targets and use a setting with 18 robots and 20 targets as our comparison baseline. The results of our experiments are shown in Figure 6. In Figure 6(a), we observe that when there are fewer (9) robots, the robots are strained to perform the tasks and most heuristics results in an average of 2-3 robots being allocated per task. This results in many tasks remaining incomplete as supported by the graph shown in Figure 4(a). When the number of robots increases while keeping the number of tasks fixed, more robots are able to perform each task (visit each target) and most tasks get completed. This behavior follows intuitively because when the number of robots increases more robots become available to perform a fixed number of tasks. An anomaly to this behavior happens for the closest task first (CT) heuristic - when the number of robots is increased to 27 the average number of robots visiting each task decreases. The inferior performance of the CT heuristic can be attributed to the fact that a robot using the CT heuristic naively selects a task based only on the distance between the robot and the task. When there are more robots in the environment, it results in all robots selecting the task closest to them, only to discover that it has been completed before most of them reach the task's location. This causes tasks that are further away to get starved and finally remain incomplete. Because of these incomplete tasks, the average number of robots visiting each task deteriorates and falls below 3. From Figure 6(b), we observe that changing the number of targets does not significantly affect the number of robots allocated to a task. For comparing the results between the different heuristics, we observe that the MPT heuristic results in more than the minimum number of robots visiting each task. This happens because the robots using the MPT heuristic are able to avoid unnecessary allocations to tasks and can efficiently complete the number of tasks in reduced time as shown in Figures 4(b) – (c).



(a) (b)

Fig. 6. Average number of robots that perform a task (visit a target) for the different aggregation heuristics used. (a) For different numbers of robots, (b) For different numbers of tasks (targets).

We have also analyzed the performance of each of the metrics reported in Figures 4-6 while varying the communication range of the robots and the speed at which the robots move.[2] We observed that the smaller communication ranges result in aggravated performance as the robots are not able to disseminate task information rapidly across all the robots within the environment. Also faster robot speeds result in slightly better performance as robots are able to visit task locations more rapidly.

6 Conclusions and Future Work

In this chapter, we described the MRTA problem for a cooperative foraging scenario and discussed different local heuristics that can be used by a robot's distributed MRTA algorithm for selecting tasks. We presented empirical results analyzing the performance of the distributed MRTA algorithm for the different heuristics in terms of the efficiency of performing different numbers of tasks in the environment. Our results show that the heuristics-based MRTA techniques provide a simple, scalable and efficient mechanism for allocating dynamically arriving tasks across multiple robots. Currently, we are investigating several directions of extending this research. We are developing improved algorithms for MRTA using

---

[2] The results of these experiments are not reported here for conciseness.

market-based and graph-based techniques. We are also adapting MRTA algorithms for use on heterogeneous robots with different types of sensors and different capabilities. Finally, we are developing mechanisms to identify and maintain robot-teams and exploring techniques to perform intra- and inter-team MRTA. Multi-robot task allocation is a crucial aspect in the efficient performance of robotic systems, and, we envisage that further research along the techniques described in this chapter will enable better understanding and improved solutions to the MRTA problem.

Acknowledgements

This research has been partially funded by the U.S. Department of Defense, Office of Naval Research through the COMRADES project, grant no. N000140911174. The author would like to thank his former students: Matthew Hoeing, David Miller, and Timothy Judkins for implementing and running some of the experiments reported in this work.

References

1. Agassounon, W., Martinoli, A.: Efficiency and Robustness of Threshold-based Distributed Allocation Algorithms in Multi-agent Systems. In: Proc. Intl. Conf. on Autonomous Agents and Multi-agent Systems (AAMAS), Bologna, Italy, pp. 1090–1097 (2002)
2. Alami, R., Fleury, S., Herrb, M., Ingrand, F., Robert, F.: Multi-robot cooperation in the MARTHA project. IEEE Robotics and Automation 5(1), 36–47 (1998)
3. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, Oxford (1999)
4. Bererton, C., Gordon, G., Thrun, S.: Auction Mechanism Design for Multi-Robot. In: Advances in Neural Information Processing Systems [Neural Information Processing Systems, NIPS 2003], vol. 16. MIT Press, Cambridge (2004)
5. Botelho, S., Alami, A.: M+: A Scheme for Multi-Robot Cooperation through Negotiated Task Allocation and Achievement. In: Proc. Intl. Conf. on Robotics and Automation, Detroit, MI, pp. 1234–1239 (1999)
6. Campos, M., Bonaeau, E., Theraulaz, G., Deneubourg, J.: Dynamic Scheduling and division of labor in social insects. Adaptive Behavior 8(2), 83–92 (2001)
7. Cheng, K., Dasgupta, P.: Dynamic Area Coverage using Faulty Multi-agent Swarms. In: Proc. IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT), Fremont, CA, pp. 17–24 (2007)
8. Cheng, K., Dasgupta, P., Wang, Y.: Distributed Area Coverage Using Robot Flocks. In: World Congress on Nature and Biologically Inspired Computing (NaBIC), pp. 678–683 (2009)
9. Correll, N., Martinoli, A.: Robust Distributed Coverage using a Swarm of Miniature Robots. In: Proc. Intl. Conf. of Robotics and Automation (ICRA), Rome, Italy, pp. 379–384 (2007)

10. Dasgupta, P.: A Multi-agent Swarming System for Distributed Automatic Target Recognition. IEEE Transactions on Systems, Man, Cybernetics, Part A 38, 549–563 (2007)
11. Dias, M.B.: Traderbots: A new paradigm for robust and efficient multi-robot coordination in dynamic environments. Ph.D. Dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (2004)
12. Gerkey, B.: On multi-robot task allocation. Ph.D Thesis, Univ. of Southern California (2003)
13. Golfarelli, M., Rizzi, S.: Spatio-temporal clustering of tasks for swap-based negotiation protocols in multi-agent systems. In: Proc. 6th International Conference on Intelligent Autonomous Systems, Venice, Italy, pp. 172–179 (2000)
14. Gurfil, P.: Evaluating UAV flock mission performance using dedek's taxonomy. In: Proc. American Conrol Conf., Portland, OR, pp. 4679–4684 (2005)
15. Hoeing, M., Dasgupta, P.: Dynamic Pricing Algorithms for Market based Distributed Task Selection in Multi-agent Swarms. In: Proc. Intl. Conf. on Intelligent Agent Technology (IAT 2006), Hong Kong, pp. 113–116 (2006)
16. Jones, E., Dias, M., Stentz, A.: Learning-enhanced market-based task allocation for disaster response. CMU Technical Report No. CMU-RI-TR-06-48 (2006)
17. Kalra, N., Martinoli, A.: A Comparative Study of Market-Based and Threshold-Based Task Allocation. In: Proc. 8th Intl. Symp. on Distributed Autonomous Robotic Systems, DARS 2006 (2006)
18. Lagoudakis, M., Berhault, M., Koenig, S., Keskinocak, P., Kleywegt, A.: Simple auctions with performance guarantees for multi-robot task allocation. In: Proc. Intl. Conf. Intelligent Robots and Systems (IROS 2004), Sandei, Japan, pp. 698–705 (2004)
19. Lagoudakis, M., Markakis, V., Kempe, D., Keskinocak, P., Koenig, S., Kleywegt, A., Tovey, C., Meyerson, A., Jain, S.: Auction-Based Multi-Robot Routing. In: Proc. of the Intl. Conf. on Robotics: Science and Systems, Boston, MA, pp. 343–350 (2005)
20. Lemaire, T., Alami, R., Lacroix, S.: A distributed tasks allocation scheme in multi-uav context. In: IEEE Intl. Conf. on Robotics and Automation, New Orleans, LA, pp. 3622–3627 (2004)
21. Lerman, K., Martinoli, A., Galstyan, A.: A Review of Probabilistic Macroscopic Models for Swarm Robotic Systems. In: Şahin, E., Spears, W.M. (eds.) Swarm Robotics 2004. LNCS, vol. 3342, pp. 143–152. Springer, Heidelberg (2005)
22. Mataric, M., Sukhatme, G., Ostergaard, E.: Multi-robot task allocation in uncertain environments. Autonomous Robots 14, 255–263 (2003)
23. Michel, O.: Webots: Professional Mobile Robot Simulation. Intl. Jour. of Advanced Robotic Systems 1(1), 39–42 (2004)
24. Miller, D., Dasgupta, P., Judkins, T.: Distributed Task Selection in Multi-agent based Swarms using Heuristic Strategies. In: Şahin, E., Spears, W.M., Winfield, A.F.T. (eds.) SAB 2006 Ws 2007. LNCS, vol. 4433, pp. 158–172. Springer, Heidelberg (2007)
25. Mondada, F., Pettinaro, G., Guignard, A., Kwee, I., Floreano, D., Deneubourg, J., Nolfi, S., Gambardella, L., Dorigo, M.: Swarm-Bot: a New Distributed Robotic Concept. Autonomous Robots 17(2-3), 193–221 (2004)
26. Nanjanath, M., Gini, M.: Dynamic Task Allocation for Robots via Auctions. In: Proc. Intl. Conf. on Robotics and Automation (ICRA 2006), Orlando, FL, pp. 2781–2786 (2006)
27. Oritz, C., Vincent, R., Morriset, B.: Task Inference and Distributed Task Management in the Centibots Robotic System. In: Proc. AAMAS 2004, Utrecht, Netherlands, pp. 870–877 (2005)

28. Parker, L.: Distributed Algorithms for Multi-Robot Observation of Multiple Moving Targets. Autonomous Robots 12(3), 231–255 (2002)
29. Sariel, S., Balch, T.: Efficient Bids on task allocation for multi-robot exploration. In: Proc. FLAIRS Conference, pp. 116–121 (2006)
30. Sariel, S., Balch, T.: Dynamic and Distributed Allocation of Resource Constrained Project Tasks to Robots. In: Proc. Multi-Agent Robotic Systems (MARS) Workshop, Setubal, Portugal, pp. 34–43 (2006)
31. Schneider, J., Apefelbaum, D., Bagnell, D., Simmons, R.: Learning opportunity costs in multi-robot market based planners. In: Proc. Intl. Conf. on Robotics and Automation (ICRA 2005), Barcelona, Spain, pp. 1151–1156 (2005)
32. Sujit, P., Sinha, A., Ghose, D.: Multiple UAV task allocation using negotiation. In: Proc. Intl. Conf. Autonomous Agent and Multi-agent Systems (AAMAS 2006), Hakodate, Japan, pp. 471–478 (2006)
33. Smith, R.: The contract net protocol: High-level communication and control in a distributed problem solver. IEEE Transactions on Computers 29, 1104–1113 (1980)
34. Tang, F., Parker, L.: A Complete Methodology for Generating Multi-Robot Task Solutions using ASyMTRe-D and Market-Based Task Allocation. In: Proc. Intl. Conf. on Robotics and Automation (ICRA 2007), Rome, Italy, pp. 3351–3358 (2007)
35. Van Parunak, H., Brueckner, S., Odell, J.: Swarming Coordination of Multiple UAVs for Collaborative Sensing. In: Proc. 2nd AIAA 'Unmanned Unlimited' Systems, Technology and Operations Aerospace, Land and Sea Conference, Workshop and Exhibits, San Diego, CA (2003)
36. Viguria, A., Maza, I., Ollero, A.: SET: An algorithm for distributed multirobot task allocation with dynamic negotiation based on task subsets. In: Proc. Intl. Conf. on Robotics and Automation (ICRA 2007), Rome, Italy, pp. 3339–3344 (2007)
37. Werger, B., Mataric, M.: Broadcast of local eligibility for multi-target observation. In: Proc. of Intl. Symp. on Distributed Autonomous Robotic Systems (DARS 2000), Knoxville, TN (2000)
38. Weyns, D., Boucke, N., Holovet, T.: Gradient field-based task assignment in an AGV transportation system. In: Proc. Intl. Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS), Hakodate, Japan, pp. 842–849 (2006)
39. Zheng, X., Koeing, S., Tovey, C.: Improving Sequential Single-Item Auctions. In: Proc. of the IEEE Intl. Conf. on Intelligent Robots and Systems (IROS), pp. 2238–2244 (2006)
40. Zlot, R.: Complex Task Allocation for Multi-robot Teams. Ph.D Thesis, Carnegie Mellon University (October 2006)

Chapter 3
Formal-Language-Oriented Foundation of Dynamics of Human Crowds Using Topos-Theory

Vladimir G. Ivancevic and Darryn J. Reid

Land Operations Division, Defence Science & Technology Organisation
Adelaide, Australia

Abstract. In contemporary neuro-quantum studies, the following main idea has been established: motion of a human body obeys the laws of Newtonian physics, while motion of human mind can be formally described only by the laws of quantum physics. From this psycho-physical modeling perspective, a general dynamics of human crowds can be divided into two categories: a relatively simplistic crowd mechanics, concerning only with the physical motion of individual agents, aggregates and a crowd as a whole; and a complex cognitive behavior of a crowd, which incorporates motivational cognition into the crowd mechanics. In this paper we will attempt to provide a topos-theoretic foundation for both crowd mechanics and general crowd behavior. For this we will use a general formal-language framework, in which crowd mechanics is described by a propositional language, while crowd behaviour is described by a higher-order typed language. This framework naturally leads to the category of complex systems and general crowd behavioral dynamics in a general topos.

1 Introduction

In contemporary quantum-brain (or, neuro-quantum) studies, the following main idea has been established: motion of a human body obeys the laws of Newtonian physics [1, 2, 3, 4], while motion of human mind can be formally described only by the laws of quantum physics[1] [7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. From this psycho–physical modeling perspective, a general dynamics of human crowds can be divided into two categories:

---

[1] According to [5], many–body quantum field theory appears to be the only existing theoretical tool capable to explain the dynamic origin of long–range correlations, their rapid and efficient formation and dissolution, their interim stability in ground states, the multiplicity of coexisting and possibly non–interfering ground states, their degree of ordering, and their rich textures relating to sensory and motor facets of behaviors. It is historical fact that many–body quantum field theory has been devised and constructed in past decades exactly to understand features like ordered pattern formation and phase transitions in condensed matter physics that could not be understood in classical physics, similar to those in the brain.

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 21–113.
springerlink.com

1. Relatively simplistic *crowd mechanics,* concerning only with the physical motion of individual agents, aggregates and a crowd as a whole. In principle, it is describable by means of classical Lagrangian/Hamiltonian mechanics and statistical mechanics (see [18, 19, 20]).
2. Complex cognitive behavior of a crowd, which incorporates motivational cognition [21, 22] into the crowd mechanics. For simplicity, we will call it *crowd behavior.* In principle, it is describable by means of quantum-like physical theories [18, 19, 23].

In this paper we will attempt to provide a topos-theoretic foundation for both crowd mechanics and general crowd behavior. Note that *topos theory* is a rich branch of mathematics which can be approached from a variety of different viewpoints (see, e.g. [24]).[2]

The domain of validity of the 'quantum' is not restricted to the microscopic world [6]. There are macroscopic features of classically behaving systems, which cannot be explained without recourse to the quantum dynamics. This field theoretic model leads to the view of the phase transition as a condensation that is comparable to the formation of fog and rain drops from water vapor, and that might serve to model both the gamma and beta phase transitions. According to such a model, the production of activity with long–range correlation in the brain takes place through the mechanism of *spontaneous breakdown of symmetry* (SBS), which has for decades been shown to describe long-range correlation in condensed matter physics. The adoption of such a field theoretic approach enables modelling of the whole cerebral hemisphere and its hierarchy of components down to the atomic level as a fully integrated macroscopic quantum system, namely as a macroscopic system which is a quantum system not in the trivial sense that it is made, like all existing matter, by quantum components such as atoms and molecules, but in the sense that some of its macroscopic properties can best be described with recourse to quantum dynamics (see [5] and references therein).

[2] At the beginning of the 20th century, philosophers of mathematics were beginning to divide into various schools of thought, broadly distinguished by their pictures of mathematical epistemology and ontology. Three schools: *formalism* (David Hilbert and his group of collaborators, including John von Neumann), *constructivism/intuitionism* (Henri Poincaré, Hermann Weyl and L.E.J. Brouwer), and *logicism* (Gottlob Frege and Bertrand Russell), emerged at this time, partly in response to the increasingly widespread worry that mathematics as it stood, and analysis in particular, did not live up to the standards of certainty and rigor that had been taken for granted. Each school addressed the issues that came to the fore at that time, either attempting to resolve them or claiming that mathematics is not entitled to its status as our most trusted knowledge. At the middle of the century, a new mathematical theory was created by Samuel Eilenberg and Saunders Mac Lane, known as *category theory* (see next section), and it became a new contender for the natural language of mathematical thinking. In our view, *topos theory* is a 'marriage' of constructivism and category theory.

In the late 1950s Alexander Grothendieck chose the Greek word *topos* (which means "place") to define a mathematical object (usually denoted by τ) that

---

would provide a general framework for his theory of *étale cohomology* and other variants related to his algebraic geometry, philosophy of descent and Grothendieck topology. Two examples of a *Grothendieck topos* are [25]:

1. The category of *sheaves*[3] of sets on a topological space is a topos. In particular, the category of sets, **Set**,[4] is a topos, for it is the category of sheaves of sets on the one point space. This topos, denoted $p\tau$, is called the *punctual topos.*
2. Let G be a group. The category BG of G-sets, i.e., sets equipped with a left action of G, is a topos. For $G = 1$, BG $= p\tau$.

What these categories have in common is that (i) they behave very much like the category **Set** of sets, and (ii) they possess a good notion of *localization.* In order to formalize (ii), Grothendieck conceived the idea of sheaf on a site, which generalizes the notion of sheaf on a topological space. That led him to the notion of topos, which encompasses (i) and (ii). The theory of a Grothendieck topos was a category of *sheaves*, where the word sheaf had acquired an extended meaning with respect to the idea of *Grothendieck topology.* However, there are logically interesting examples of topoi that are not Grothendieck topoi (or, geometrical topoi).

[3] Sheaf theory is a subfield of topology that generalizes the notion of gluing continuous functions together and specializes in describing objects by their local data on open sets. Above all, it's a tool, usable to define connectedness, manifolds, the fundamental group, and several other concepts in algebraic topology and differential geometry. More specifically, a *sheaf* is a tool for systematically tracking locally defined data attached to the open sets of a topological space. The data can be restricted to smaller open sets, and the data assigned to an open set is equivalent to all collections of compatible data assigned to collections of smaller open sets covering the original one. For example, such data can consist of the rings of continuous or smooth real-valued functions defined on each open set. The first step in defining a sheaf is to define a presheaf, which captures the idea of associating data and restriction maps to the open sets of a topological space. In particular, a presheaf of sets on a topological space is a structure which associates to each open set U of the space a set $F(U)$ of 'sections' on U, and to each open set V included in U a map $F(U) \to F(V)$ giving restrictions of sections over U to V. Sheaves attach to the properties of presheaves the ability to paste together sections on different domains. So, the second step is to require the normalization and gluing axioms. A presheaf which satisfies these axioms is a sheaf. There are presheaves which are not sheaves (e.g., let X be the real line, and let $F(U)$ be the set of bounded continuous functions on U. This is not a sheaf because it is not always possible to glue). There are also morphisms (or, maps) from one sheaf to another (the definition of a morphism on sheaves makes sense only for sheaves on the same space X). That is, sheaves of a specific type (e.g., sheaves of Abelian groups) with their morphisms on a fixed topological space form a category. A morphism of sheaves is defined as a collection of functions, one for each open set, which satisfy a compatibility condition. Alternatively, a morphism of sheaves is a natural transformation of the corresponding functors (see next section). On the other hand, to each continuous map there is associated both a direct image functor, taking sheaves and their morphisms on the domain to sheaves and morphisms on the codomain, and an inverse image functor operating in the opposite direction.

[4] More precisely, *small* sets and functions between them. Small means that we do not have proper classes. One must take care in these foundational issues to avoid problems like Russell's paradox.

Current definition of topos goes back to 1963, when William Lawvere decided to formulate new foundations for mathematics, based on category theory (see [26]). Their definition picks out the central role in topos theory of the *sub-object classifier.* In the usual category of sets, this is the two-element set of Boolean truth-values, *true* and *false.* The subsets of a given set X are the same as the functions on X to any such given two-element set: fix the 'first' element and make a subset Y correspond to the function sending Y there and its complement in X to the other element. Lawvere and Tierney formulated axioms for a topos that assumed a sub-object classifier, and some limit conditions, to make a *Cartesian-closed category.* For a while this notion of topos was called *elementary topos.*

From one point of view (Lawvere), a topos is a category with certain properties characteristic of the category of sets, **Set**, a sort of a generalized set theory with some special constructions in **Set**. From another point of view (Grothendieck), a topos is an abstraction of the category of sheaves over a topological space. Briefly, a topos τ is a category which has the following two basic properties (see [27, 28, 29]):

1. All limits taken over finite index categories exist.
2. Every object has a power object.

From these two basic properties, one can derive four another properties:

(a) All colimits taken over finite index categories exist.[5]
(b) The category τ has a sub-object classifier.[6]
(c) Any two objects in the topos τ have an exponential object.[7]
(d) The category τ is Cartesian closed.[8]

---

[5] This property (together with the property 1. above) means that in the topos τ there are: (i) an initial object (an object like the empty set); (ii) a terminal object (an object like a set with one element); (iii) binary coproducts (something like the disjoint union of two sets; (iv) binary products (something like the Cartesian product of two sets); (v) equalizers (something like the subset of X consisting of all elements x such that $f(x) = g(x)$, where $f, g : X \longrightarrow Y$); and (vi) coequalizers (something like the quotient set of X, where two elements $f(y)$ and $g(y)$ are identified, where $f, g : X \longrightarrow Y$).

[6] This means that in the topos τ there is an object Ω called the 'sub-object classifier', which acts like {0,1}, in that functions from any set x into {0,1} are 'secretly' the same as subsets of x. We can think of Ω as the replacement for the usual Boolean truth values 'true' and 'false'.

[7] For any objects x and y in τ, there is an object y^x, called an 'exponential', which acts like 'the set of functions from x to y'.

[8] A category τ is Cartesian closed if any morphism defined on a Cartesian product $A \times B$ of two objects $A, B \in \tau$ can be naturally identified with a morphism defined on one of the factors A or B. Cartesian closed categories are particularly important in mathematical logic and computer science.

Because topos is a special type of category, it consists of objects, and morphisms (or, arrows) from one object to another, so that in certain critical respects it behaves like the category of sets, **Set**. In particular, the 'opposite' of a category C is a category, denoted C^{op}, whose objects are the same as those of C, and whose morphisms are defined to be the opposite of those of C; i.e., a morphism $f : A \to B$ in C^{op} is said to exist if and only if (iff, for short) there is a morphism $f : B \to A$ in C. Also, an object 1 is said to be a terminal object in a category C if there is just one morphism from any other object in C to 1; it is easy to see that any two terminal objects in C are isomorphic. In the category **Set**, a terminal object is any set $\{*\}$ with just a single element, so it is called a *singleton*. In this case a morphism is just a map, and hence a morphism $\{*\} \to X$ picks out a unique element of X.

For example, if C is a small category, then the associated *functor category* $\mathbf{Set}^C$ (consisting of all covariant functors from C to the category **Set**, with natural transformations as morphisms, see next section) is a topos. For instance, the category *Grph* of graphs of the kind permitting multiple directed edges between two vertices is a topos. A graph consists of two sets, an edge set and a vertex set, and two functions s, t between those sets, assigning to every edge e its source $s(e)$ and target $t(e)$. The topos *Grph* is thus equivalent to the functor category $\mathbf{Set}^C$, where C is the category with two objects E and V and two morphisms $s, t : E \longrightarrow V$ giving respectively the source and target of each edge. More generally, the categories of finite sets, of finite $G-$sets (actions of a group G on a finite set), and of finite graphs are topoi.

Another defining property for a category C to be a topos is that a product $A \times B$ exists for any pair of objects (A, B) in C. Yet another of the basic properties of a topos is that there is a $1-1$ correspondence between morphisms $f : A \times B \to \Omega$ and morphisms $\ulcorner f \urcorner : A \to PB := \Omega^B$, where is the so-called 'sub-object classifier' (explained later). In general, $\ulcorner f \urcorner$ is called the *power transpose* of f. If $A \simeq 1$ then $\ulcorner f \urcorner$ is known as the *name* of the morphism $f : B \to \Omega$.

From logical perspective, topos-theoretic approach leads to a picture in which the 'truth values', or 'semantic values' of such contextual predictions are not just two-valued (i.e., 'true' and 'false') but instead lie in a larger logical algebra, the so–called *Heyting algebra.* Named after Dutch mathematician Arend Heyting (a student of L. Brouwer), the Heyting algebras are algebraic structures that play in relation to *intuitionistic logic*[9]

---

[9] The name 'intuitionistic' comes from its original motivation of a formal basis for *Brouwer's intuitionism.* The name 'constructive logic' would be a better name to use, but most references stick with 'intuitionistic', and thus so will we.

a role analogous to that played by *Boolean algebras* in relation to *classical logic*.[10] A Heyting algebra is a pseudocomplemented, distributive

---

[10] Recall that classical Aristotelian logic has only two truth-values, 'true' and 'false' (usually denoted by T, F, or 1,0), and includes five propositional truth functions: (1) negation, or logical *not*: $(\neg a)$; (2) conjunction, or logical *and*: $(a \wedge b)$; (3) disjunction, or logical *or*: $(a \vee b)$; (4) material implication, or logical *if-then*: $(a \Rightarrow b)$; and (5) material equivalence, or logical *if-and-only-if* (or, *iff*, for short): $(a \Leftrightarrow b)$. In particular, both $(a \Rightarrow \neg\neg a)$ and $(\neg\neg a \Rightarrow a)$ are *theorems*, while conjunctions and disjunctions are combined into De Morgan's laws: $(\neg(a \vee b) = \neg a \wedge \neg b)$; and $(\neg(a \wedge b) = \neg a \vee \neg b)$. An expression is a *tautology* if it is true for all variable assignments. For example, the following expression is tautology: $t = (p \vee \neg r) \wedge s \Rightarrow (r \wedge s \Rightarrow (p \wedge q) \vee p)$. An expression is a *contradiction* if it is false for all variable assignments.

The basic *inference rules* of Aristotelian syllogistic logic are:

Modus ponens: $((a \Rightarrow b) \wedge a) \Rightarrow b$;
Modus tollens: $((a \Rightarrow b) \wedge \neg b) \Rightarrow (\neg a)$;
Modus tollendo ponens: $((a \vee b) \wedge \neg a) \Rightarrow b$;
Modus ponendo tollens: $(\neg(a \wedge b) \wedge a) \Rightarrow (\neg b)$; and
Reductio ad absurdum $((a \Rightarrow b) \wedge (a \Rightarrow (\neg b))) \Rightarrow (\neg a)$.

On the other hand, *intuitionistic logic*, or *constructive logic*, can be described as classical logic without the *Aristotelian law of excluded middle*: $(a \vee \neg a)$(since one can construct, via Gödel's *incompleteness theorems*, a mathematical statement that can be neither proven nor disproved); as well as without the *elimination of double negation*: $(\neg\neg a \Rightarrow a)$; and without Peirce's law: $((a \Rightarrow b) \Rightarrow a) \Rightarrow a$; but with the law of contradiction: $(\neg a \Rightarrow (a \Rightarrow b))$, and with Modus ponens as a main inference rule. Intuitionistic logic encompasses the principles of logical reasoning which were used by L.E.J. Brouwer in developing his intuitionistic mathematics, beginning in 1907.

Philosophically, intuitionism differs from logicism by treating logic as a part of mathematics rather than as the foundation of mathematics. It differs from finitism by allowing (constructive) reasoning about infinite collections. Finally, it differs from Platonism by viewing mathematical objects as mental constructs with no independent ideal existence. This last difference between intuitionism and Platonism seems to be the deepest, which is consistent with metamathematics' view including algorithmic information theory, theory of computation and the foundations of computer science. Note, though, that it does not mean Aristotleism either.

A fundamental fact about intuitionistic logic is that it has the same consistency strength as classical logic. For propositional logic this was first proved by Glivenko's Theorem 1929: An arbitrary propositional formula a is classically provable, iff $(\neg\neg a)$ is intuitionistically provable. Kurt Gödel proved in 1933 the *equiconsistency* of intuitionistic and classical theories [31]. In particular, Gödel observed that intuitionistic propositional logic has the disjunction property: if $(a \vee b)$ is a theorem, then a is a theorem or b is a theorem. An embedding of classical first-order logic into intuitionistic logic is given by the Gödel–Gentzen *double-negation translation*.

lattice[11] with zero element 0 and unit element 1, representing 'totally false' resp. 'totally true'. The pseudocomplement is denoted by $\neg$, and one has, for all elements α of a Heyting algebra H,

$$\alpha \vee \neg\alpha \leq 1,$$

in contrast to $\alpha\vee\neg\alpha = 1$ in a Boolean algebra. This means that the disjunction of a proposition α and its negation need not be (totally) true in a Heyting algebra. Equivalently, one has

$$\neg\neg\alpha \geq \alpha,$$

in contrast to $\neg\neg\alpha = \alpha$ in familiar Boolean algebras. This shows that Boolean logic is a special case of intuitionistic logic.

Just as normal set theory is intimately associated with Boolean algebra (i.e., the 'Venn diagram' algebra of subsets of a set is Boolean), so a topos is associated with a more general Heyting algebra connected to the sub-objects of objects in the topos. In particular, *fuzzy set theory* (see, e.g. [17]) can be viewed as a sub-branch of topos theory.

We start our general crowd modeling with the following basic assumption: Constructing a general crowd dynamics is equivalent to finding a representation in a topos of a certain *formal language,* either a (relatively simple) *propositional language* or a (more sophisticated) *higher-order typed language*, that is attached to the crowd system. Crowd mechanics arises when the topos is the category of sets, **Set**. Other types of crowd behavior employ a different, more general topos.

The other important aspect of intuitionistic reasoning is that here we essentially treat application of axiom schemes or inference rules as *resources* – and hence its significance in representing 'computational' things such as crowds.

Now we turn to the physical side of the story. Based on the previous work of C. Isham [34, 35, 36, 37], in a recent series of four papers in J. Math. Phys., A. Döring and Isham proposed a topos-theoretic foundation for various classical and quantum theories of physics, with quantum-gravity as a main goal [38, 39, 40, 41]. As we are going to apply *Döring-Isham approach* to our general crowd dynamics, we give here a brief review of their four papers.

In the first paper [38], they introduced the idea that a formal language could be attached to each physical system S and that, in the broadest sense, constructing a physical theory of S was equivalent to finding a representation ϕ of this language in a topos τ_ϕ. Constructing a theory of physics was then

---

[11] Lattice is meant in the algebraic sense: a partially ordered set L such that any two elements $a, b \in L$ have a minimum (greatest lower bound) $a \wedge b$ and a maximum (least upper bound) $a \vee b$ in L. A lattice L is distributive if and only if $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ as well as $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ hold for all $a, b, c \in L$.

equivalent to finding a representation of this language in a topos. Two different kinds of language were discussed: a simple propositional language $\mathcal{PL}(S)$; and a more sophisticated, higher-order, typed language (a 'local' language) $\mathcal{TL}(S)$. The language $\mathcal{PL}(S)$ provides a deductive system (using intuitionistic logic) and hence provides a way of making statements about the physical system S. However, a purely propositional language is limited in scope: at the very least, one would like to have a 'first-order' language, so that the phrases 'for all' and 'there exists' can be used. Besides, such a language is rudimentary in so far as many features of a physical theory would lie *outside* its scope, and are introduced only when constructing a representation. For example, in classical mechanics, the entities that lie outside the language are (i) the state space $\mathcal{S}$; (ii) the choice of $\mathbb{R}$ as the set in which physical quantities take their values; (iii) the specific subset $\varDelta \subseteq \mathbb{R}$ that is used in the proposition '$A \in \varDelta$' and (iv) the real-valued functions on $\mathcal{S}$ that represent physical quantities. For this reason, the next step was to assign to each physical system S, a more powerful, typed language, $\mathcal{TL}(S)$. Their general scheme could then be understood as the task of finding representations of $\mathcal{TL}(S)$ in various topoi. The language $\mathcal{TL}(S)$ has two 'ground-type' symbols, $\varSigma$ and $\mathcal{R}$, and a set of 'function symbols', written rather suggestively as the string of characters '$A : \varSigma \to \mathcal{R}$'. These are the linguistic precursors of, respectively, the state object, the quantity-value object,[12] and the morphisms between them that represent physical quantities. A symbol $\tilde{\varDelta}$ could be introduced as a variable of type $P\mathcal{R}$. By these means, the entities that lie outside the propositional language $\mathcal{PL}(S)$ are all brought 'inside' the local language $\mathcal{TL}(S)$.

The second paper [39], dealt with the topos representation of the propositional language $\mathcal{PL}(S)$ in quantum theory. The motivation came from a desire to address certain deep issues that arise when contemplating quantum theories of space and time. They studied in depth the topos representation of the propositional language, $\mathcal{PL}(S)$, for the case of quantum theory. In doing so, they made a direct link with, and clarify, their earlier work on applying topos theory to quantum physics. The key step was a process they term 'daseinisation' after by which a projection operator is mapped to a sub-object of the spectral presheaf—the topos quantum analogue of a classical state space.

For the language $\mathcal{PL}(S)$, a key result from the topos constructions is that, given any quantum state $|\,\psi\,\rangle$, there are generalized truth values, $\nu\big(A \in \varDelta; |\,\psi\,\rangle\big)$, for propositions '$A \in \varDelta$'. These truth values belong to the Heyting algebra $\varGamma\varOmega_{\phi}$ of global elements of the sub-object classifier $\varOmega_{\phi}$ of the topos concerned. In making these assignments, nothing is said about 'measurements', or 'observers', or even 'probability': there is just the truth value $\nu\big(A \in \varDelta; |\,\psi\,\rangle\big) \in \varGamma\varOmega_{\phi}$.

In the second part of [39] the authors changed gear with the introduction of the more sophisticated local language $\mathcal{TL}(S)$. From this point forward,

---

[12] These foundational terms of *type theory* are applicable to programming languages (among other things).

throughout the rest of the series of papers, their attention has been devoted almost entirely to this language. More specifically, they used $\mathcal{TL}(S)$ to study 'truth objects' in the topos. These are objects in the topos that play the role of states: a necessary development as the spectral presheaf has no global elements, and hence there are no microstates in the sense of classical physics.

In the third paper [40], Döring and Isham returned to the language $\mathcal{TL}(S)$ and tried to find a topos representation ϕ for quantum theory. The topos of the $\mathcal{PL}(S)$-representation is $\mathbf{Set}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$: the topos of presheaves over the category, $\mathcal{V}(\mathcal{H})$, of unital, Abelian subalgebras of the algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators on the quantum Hilbert space $\mathcal{H}$. They used the same topos for the $\mathcal{TL}(S)$-representation, with the spectral presheaf, $\underline{\Sigma}$, being identified as the $\mathbf{Set}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$-representative, Σ_ϕ, of the ground-type symbol Σ. Thus $\underline{\Sigma}$ is the state object, and, therefore, propositions are represented by sub-objects of $\underline{\Sigma}$; just as in classical physics, a proposition about the system is represented by a subset of the classical state space. The steps in finding the representation of $\mathcal{TL}(S)$ are first to identify the quantity-value object, $\mathcal{R}_\phi$; and then to show how to represent a physical quantity, A, by a morphism $\breve{\delta}(A) : \underline{\Sigma} \to \mathcal{R}_\phi$. Both problems were solved in the third paper.

The motivation for the last paper in the series [41], came from the authors' desire to address certain deep issues that arise in the quantum theory of gravity. Their basic contention was that constructing a theory of physics was equivalent to finding a representation in a topos of a certain formal language that was attached to the system. Classical physics arises when the topos is the category of sets. Other types of theory employ a different topos. More specifically, they turned to considering a *collection* of systems: in particular, they were interested in the relation between the topos representation for a composite system, and the representations for its constituents. They also studied that problem for the disjoint sum of two systems. Their approach to these matters was to construct a *category* of physical systems and to find a topos representation of the entire category.

In the present paper, we will apply Döring–Isham foundational topos theory to general crowd dynamics, so that their 'classical physics' translates into our 'crowd mechanics', while their 'quantum physics' translates into our 'crowd behavior':

$$\begin{aligned} DI : \text{classical physics} \quad &\Longrightarrow \quad IR : \text{crowd mechanics}, \\ DI : \text{quantum physics} \quad &\Longrightarrow \quad IR : \text{crowd behavior}. \end{aligned}$$

The plan of the paper is as follows. Firstly, we will give a brief overview of category theory, followed by a brief review of topos theory, including only details that are relevant for our crowd modelling. Then, we will describe the role of formal languages in crowd modelling, first a propositional language $\mathcal{PL}(S)$, and then a higher-order, typed language $\mathcal{TL}(S)$. After that, we will describe our general category of complex systems. We will finish with a general crowd dynamics in a general topos.

2 A Brief on Category Theory

In modern mathematical sciences whenever one defines a new class of mathematical objects, one proceeds almost in the next breath to say what kinds of maps between objects will be considered [58, 52, 59, 3, 4]. A general framework for dealing with situations where we have some *objects* and *maps between objects*, like sets and functions, vector spaces and linear operators, points in a space and paths between points, etc. – gives the modern metalanguage of categories and functors. Categories are mathematical universes and functors are 'projectors' from one universe onto another.

2.1 Maps

The reader is assumed to be familiar with usual notations of *set theory* such as $\in, \cup, \cap$. In this section we will rather emphasize the notion of a *map* between two sets, while in the next section we will introduce the notion of a *topological space*, which is basically a set with a bit of an additional structure (topology) defined on it.

We recall that a map (or, a *function*) f is a *rule* that assigns to each element x in a set A exactly one element, called $f(x)$, in a set B. A map could be thought of as a *machine* $[[f]]$ with $x-$input (the *domain* of f is the set of all possible inputs) and $f(x)-$output (the *range* of f is the set of all possible outputs) (see [60])

$$x \to [[f]] \to f(x).$$

There are four possible ways to represent a function (or, a map): (i) verbally (by a description in words); (ii) numerically (by a table of values); (iii) visually (by a graph); and (iv) algebraically (by an explicit formula). The most common method for visualizing a function is its *graph*.[13] If f is a function with domain A, then its graph is the set of ordered input–output pairs

$$\{(x, f(x)) : x \in A\}.$$

Given a map (or, a function) $f : A \to B$, the set A is called the *domain* of f, and denoted $\mathrm{Dom}\, f$. The set B is called the *codomain* of f, and denoted $\mathrm{Cod}\, f$. The codomain is not to be confused with the *range* of $f(A)$, which is in general only a subset of B.[14]

---

[13] A generalization of the graph concept is a concept of a *cross–section of a fibre bundle*, which is one of the core geometrical objects for dynamics of complex systems (see [3]).

[14] However, from a formal language point of view, we may algebraically alow higher-order functions, which cannot be represented in terms of simple mappings between sets of atoms. *Lambda calculi,* for instance, allow functions that map between functions, which places the domain and range directly inside the function rather than being external to it. The sophistication of this meant that even basic lambda calculus, despite dating to Church's work in 1920, did not have a model theory until the 60s.

A map $f : X \to Y$ is called *injective*, or 1–1, or an *injection*, iff for every y in the codomain Y there is *at most* one x in the domain X with $f(x) = y$. Put another way, given x and x' in X, if $f(x) = f(x')$, then it follows that $x = x'$. A map $f : X \to Y$ is called *surjective*, or *onto*, or a *surjection*, iff for every y in the codomain $\operatorname{Cod} f$ there is *at least* one x in the *domain* X with $f(x) = y$. Put another way, the *range* $f(X)$ is equal to the codomain Y. A map is *bijective* iff it is both injective and surjective. Injective functions are called *monomorphisms*, and surjective functions are called *epimorphisms* in the *category of sets* (see below). Bijective functions are called *isomorphisms*.

A *relation* is any subset of a *Cartesian product* of two sets (explained below). By definition, an *equivalence relation* α on a set X is a relation which is *reflexive, symmetrical* and *transitive*, i.e., relation that satisfies the following three conditions:

1. *Reflexivity*: each element $x \in X$ is equivalent to itself, i.e., $x\alpha x$;
2. *Symmetry*: for any two elements $a, b \in X$, $a\alpha b$ implies $b\alpha a$; and
3. *Transitivity*: $a\alpha b$ and $b\alpha c$ implies $a\alpha c$.

Similarly, a relation $\leq$ defines a *partial order* on a set S if it has the following properties:

1. *Reflexivity*: $a \leq a$ for all $a \in S$;
2. *Antisymmetry*: $a \leq b$ and $b \leq a$ implies $a = b$; and
3. *Transitivity*: $a \leq b$ and $b \leq c$ implies $a \leq c$.

A *partially ordered set* (or *poset*) is a set taken together with a partial order on it. Formally, a partially ordered set is defined as an ordered pair $P = (X, \leq)$, where X is called the *ground set* of P and $\leq$ is the partial order of P.

Let f and g be maps with domains A and B. Then the maps $f + g$, $f - g$, fg, and f/g are defined as follows [3]

$$\begin{aligned}
(f+g)(x) &= f(x) + g(x) && \text{domain } = A \cap B, \\
(f-g)(x) &= f(x) - g(x) && \text{domain } = A \cap B, \\
(fg)(x) &= f(x)\, g(x) && \text{domain } = A \cap B, \\
\left(\frac{f}{g}\right)(x) &= \frac{f(x)}{g(x)} && \text{domain } = \{x \in A \cap B : g(x) \neq 0\}.
\end{aligned}$$

Given two maps f and g, the composite map $f \circ g$, called the *composition* of f and g,[15] is defined by

$$(f \circ g)(x) = f(g(x)).$$

The $(f \circ g)-$machine is composed of the $g-$machine (first) and then the $f-$machine [60],

$$x \to [[g]] \to g(x) \to [[f]] \to f(g(x)).$$

For example, suppose that $y = f(u) = \sqrt{u}$ and $u = g(x) = x^2 + 1$. Since y is a function of u and u is a function of x, it follows that y is ultimately a function of x. We calculate this by substitution

$$y = f(u) = f \circ g = f(g(x)) = f(x^2 + 1) = \sqrt{x^2 + 1}.$$

2.2 Topological Spaces

Topology is a kind of *abstraction* of Euclidean geometry, and also a natural framework for the study of *continuity.*[16] Euclidean geometry is abstracted by regarding triangles, circles, and squares as being the same basic object. Continuity enters because in saying this one has in mind a *continuous deformation* of a triangle into a square or a circle, or any arbitrary shape. On the other hand, a disk with a hole in the center is topologically different from a

---

[15] If f and g are both differentiable (or smooth, i.e., C^∞) maps and $h = f \circ g$ is the composite map defined by $h(x) = f(g(x))$, then h is differentiable and h' is given by the product [60]

$$h'(x) = f'(g(x))\, g'(x).$$

In Leibniz notation, if $y = f(u)$ and $u = g(x)$ are both differentiable maps, then

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

The reason for the name *chain rule* becomes clear if we add another link to the chain. Suppose that we have one more differentiable map $x = h(t)$. Then, to calculate the derivative of y with respect to t, we use the chain rule twice,

$$\frac{dy}{dt} = \frac{dy}{du}\frac{du}{dx}\frac{dx}{dt}.$$

[16] Intuitively speaking, a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous near a point x in its domain if its value does not jump there. That is, if we just take δx to be small enough, the two function values $f(x)$ and $f(x + \delta x)$ should approach each other arbitrarily closely. In more rigorous terms, this leads to the following definition: A function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous at $x \in \mathbb{R}$ if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $y \in \mathbb{R}$ with $|y - x| < \delta$, we have that $|f(y) - f(x)| < \epsilon$. The whole function is called continuous if it is continuous at every point x.

circle or a square because one cannot create or destroy holes by continuous deformations. Thus using topological methods one does not expect to be able to identify a geometrical figure as being a triangle or a square. However, one does expect to be able to detect the presence of gross features such as holes or the fact that the figure is made up of two disjoint pieces etc. In this way topology produces theorems that are usually qualitative in nature – they may assert, for example, the existence or non–existence of an object. They will not, in general, give the means for its construction [61].

Study of topology starts with the fundamental notion of *topological space.* Let X be any *set* and $Y = \{X_\alpha\}$ denote a collection, finite or infinite of subsets of X. Then X and Y form a topological space provided the X_α and Y satisfy:

1. Any finite or infinite subcollection $\{Z_\alpha\} \subset X_\alpha$ has the property that $\cup Z_\alpha \in Y$;
2. Any *finite subcollection* $\{Z_{\alpha_1}, ..., Z_{\alpha_n}\} \subset X_\alpha$ has the property that $\cap Z_{\alpha_i} \in Y$; and
3. Both X and the empty set belong to Y.

The set X is then called a topological space and the X_α are called *open sets.* The choice of Y satisfying (2) is said to give a topology to X.

Given two topological spaces X and Y, a map $f : X \to Y$ is *continuous* if the inverse image of an open set in Y is an open set in X.

The main general idea in topology is to study spaces which can be continuously deformed into one another, namely the idea of *homeomorphism.* If we have two topological spaces X and Y, then a map $f : X \to Y$ is called a homeomorphism iff

1. f is continuous (C^0), and
2. There exists an inverse of f, denoted f^{-1}, which is also continuous.

Definition (2) implies that if f is a homeomorphism then so is f^{-1}. Homeomorphism is the main topological example of *reflexive*, *symmetrical* and *transitive relation*, i.e., *equivalence relation.* Homeomorphism divides all topological spaces up into *equivalence classes.* In other words, a pair of topological spaces, X and Y, belong to the same equivalence class if they are homeomorphic.

The second example of topological equivalence relation is *homotopy.* While homeomorphism generates equivalence classes whose members are topological spaces, homotopy generates equivalence classes whose members are continuous (C^0) maps. Consider two continuous maps $f, g : X \to Y$ between topological spaces X and Y. Then the map f is said to be *homotopic* to the map g if f can be continuously deformed into g (see below for the precise definition of homotopy). Homotopy is an equivalence relation which divides the space of continuous maps between two topological spaces into equivalence classes [61].

Another important notions in topology are *covering*, *compactness* and *connectedness.* Given a family of sets $\{X_\alpha\} = X$ say, then X is a *covering* of another set Y if $\cup X_\alpha$ contains Y. If all the X_α happen to be open sets the covering is called an *open covering.* Now consider the set Y and all its possible open coverings. The set Y is *compact* if for every open covering $\{X_\alpha\}$ with $\cup X_\alpha \supset Y$ there always exists a finite subcovering $\{X_1, ..., X_n\}$ of Y with $X_1 \cup ... \cup X_n \supset Y$. Again, we define a set Z to be *connected* if it cannot be written as $Z = Z_1 \cup Z_2$, where Z_1 and Z_2 are both open non–empty sets and $Z_1 \cap Z_2$ is an empty set.

Let $A_1, A_2, ..., A_n$ be closed subspaces of a topological space X such that $X = \cup_{i=1}^n A_i$. Suppose $f_i : A_i \to Y$ is a function, $1 \leq i \leq n$, such that

$$f_i|A_i \cap A_j = f_j|A_i \cap A_j, \qquad (1 \leq i, j \leq n). \tag{1}$$

In this case f is continuous iff each f_i is. Using this procedure we can define a C^0-function $f : X \to Y$ by cutting up the space X into closed subsets A_i and defining f on each A_i separately in such a way that $f|A_i$ is obviously continuous; we then have only to check that the different definitions agree on the *overlaps* $A_i \cap A_j$.

The *universal property of the Cartesian product*: let $p_X : X \times Y \to X$, and $p_Y : X \times Y \to Y$ be the *projections* onto the first and second factors, respectively. Given any pair of functions $f : Z \to X$ and $g : Z \to Y$ there is a unique function $h : Z \to X \times Y$ such that $p_X \circ h = f$, and $p_Y \circ h = g$. Function h is continuous iff both f and g are. This property characterizes $X \times Y$ up to isomorphism. In particular, to check that a given function $h : Z \to X$ is continuous it will suffice to check that $p_X \circ h$ and $p_Y \circ h$ are continuous.

The *universal property of the quotient*: let α be an equivalence relation on a topological space X, let X/α denote the *space of equivalence classes* and $p_\alpha : X \to X/\alpha$ the *natural projection.* Given a function $f : X \to Y$, there is a function $f' : X/\alpha \to Y$ with $f' \circ p_\alpha = f$ iff $x\alpha x'$ implies $f(x) = f(x')$, for all $x \in X$. In this case f' is continuous iff f is. This property characterizes X/α up to homeomorphism.

Homotopy. Now we return to the fundamental notion of homotopy. Let I be a compact unit interval $I = [0, 1]$. A *homotopy* from X to Y is a continuous function $F : X \times I \to Y$. For each $t \in I$ one has $F_t : X \to Y$ defined by $F_t(x) = F(x, t)$ for all $x \in X$. The functions F_t are called the 'stages' of the homotopy. If $f, g : X \to Y$ are two continuous maps, we say f is homotopic to g, and write $f \simeq g$, if there is a homotopy $F : X \times I \to Y$ such that $F_0 = f$ and $F_1 = g$. In other words, f can be continuously deformed into g through the stages F_t. If $A \subset X$ is a subspace, then F is a homotopy relative to A if $F(a, t) = F(a, 0)$, for all $a \in A, t \in I$.

The homotopy relation $\simeq$ is an equivalence relation. To prove that we have $f \simeq f$ is obvious; take $F(x, t = f(x)$, for all $x \in X$, $t \in I$. If $f \simeq g$ and F is a homotopy from f to g, then $G : X \times I \to Y$ defined by $G(x, t) = F(x, 1-t)$, is a homotopy from g to f, i.e., $g \simeq f$. If $f \simeq g$ with homotopy F and $g \simeq f$ with homotopy G, then $f \simeq h$ with homotopy H defined by

$$H(x,t) = \begin{cases} F(x,t), & 0 \leq t \leq 1/2 \\ G(x, 2t-1), & 1/2 \leq t \leq 1 \end{cases}.$$

To show that H is continuous we use the relation (1).

In this way, the set of all C^0-functions $f : X \to Y$ between two topological spaces X and Y, called the *function space* and denoted by Y^X, is partitioned into equivalence classes under the relation $\simeq$. The equivalence classes are called *homotopy classes*, the homotopy class of f is denoted by $[f]$, and the set of all homotopy classes is denoted by $[X; Y]$.

If α is an equivalence relation on a topological space X and $F : X \times I \to Y$ is a homotopy such that each stage F_t factors through X/α, i.e., $x\alpha x'$ implies $F_t(x) = F_t(x')$, then F induces a homotopy $F' : (X/\alpha) \times I \to Y$ such that $F' \circ (p_\alpha \times 1) = F$.

Homotopy theory has a range of applications of its own, outside topology and geometry, as for example in proving Cauchy theorem in complex variable theory, or in solving nonlinear equations of artificial neural networks.

A *pointed set* (S, s_0) is a set S together with a distinguished point $s_0 \in S$. Similarly, a *pointed topological space* (X, x_0) is a space X together with a distinguished point $x_0 \in X$. When we are concerned with pointed spaces $(X, x_0), (Y, y_0)$, etc, we always require that all functions $f : X \to Y$ shell preserve base points, i.e., $f(x_0) = y_0$, and that all homotopies $F : X \times I \to Y$ be relative to the base point, i.e., $F(x_0, t) = y_0$, for all $t \in I$. We denote the homotopy classes of base point–preserving functions by $[X, x_0; Y, y_0]$ (where homotopies are relative to x_0). $[X, x_0; Y, y_0]$ is a pointed set with base point f_0, the constant function: $f_0(x) = y_0$, for all $x \in X$.

A *path* $\gamma(t)$ from x_0 to x_1 in a topological space X is a continuous map $\gamma : I \to X$ with $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Thus X^I is the space of all paths in X with the compact–open topology. We introduce a relation $\sim$ on X by saying $x_0 \sim x_1$ iff there is a path $\gamma : I \to X$ from x_0 to x_1. Clearly, $\sim$ is an equivalence relation; the set of equivalence classes is denoted by $\pi_0(X)$. The elements of $\pi_0(X)$ are called the *path components*, or $0-$*components* of X. If $\pi_0(X)$ contains just one element, then X is called *path connected*, or $0-$*connected*. A *closed path*, or *loop* in X at the point x_0 is a path $\gamma(t)$ for which $\gamma(0) = \gamma(1) = x_0$. The *inverse loop* $\gamma^{-1}(t)$ based at $x_0 \in X$ is defined by $\gamma^{-1}(t) = \gamma(1-t)$, for $0 \leq t \leq 1$. The *homotopy of loops* is the particular case of the above defined homotopy of continuous maps.

If (X, x_0) is a pointed space, then we may regard $\pi_0(X)$ as a pointed set with the $0-$component of x_0 as a base point. We use the notation $\pi_0(X, x_0)$ to denote $p_0(X, x_0)$ thought of as a pointed set. If $f : X \to Y$ is a map then f sends $0-$components of X into $0-$components of Y and hence defines a function $\pi_0(f) : \pi_0(X) \to \pi_0(Y)$. Similarly, a base point–preserving map $f : (X, x_0) \to (Y, y_0)$ induces a map of pointed sets $\pi_0(f) : \pi_0(X, x_0) \to \pi_0(Y, y_0)$. In this way defined π_0 represents a ‘functor’ from the ‘category’ of topological (point) spaces to the underlying category of (point) sets (see the next subsection).

The *fundamental group* (introduced by Poincaré), denoted $\pi_1(X)$, of a pointed space (X, x_0) is the group formed by the equivalence classes of the set of all *loops*, i.e., closed homotopies with initial and final points at a given base point x_0. The identity element of this group is the set of all paths homotopic to the degenerate path consisting of the point x_0.[17] The fundamental group $\pi_1(X)$ only depends on the homotopy type of the space X, that is, fundamental groups of homeomorphic spaces are isomorphic.

Combinations of topology and calculus give differential topology and differential geometry.

2.3 Commutative Diagrams

The *category theory* (see below) was born with an observation that many properties of mathematical systems can be unified and simplified by a presentation with *commutative diagrams of morphisms* [58, 52]. Each morphism $f : X \to Y$ represents a function (i.e., a map, transformation, operator); that is, a source (domain) set X, a target (codomain) set Y, and a rule $x \mapsto f(x)$ which assigns to each element $x \in X$ an element $f(x) \in Y$. A typical diagram of sets and functions is

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[dr, "h"'] & Y \arrow[d, "g"] \\
& Z
\end{tikzcd}
\quad \text{or} \quad
\begin{tikzcd}
X \arrow[r, "f"] \arrow[dr, "h"'] & f(X) \arrow[d, "g"] \\
& g(f(X))
\end{tikzcd}$$

This diagram is *commutative* iff $h = g \circ f$, where $g \circ f$ is the usual composite function $g \circ f : X \to Z$, defined by $x \mapsto g(f(x))$.

Similar commutative diagrams apply in other mathematical, physical and computing contexts; e.g., in the 'category' of all topological spaces, the letters X, Y, and Z represent topological spaces while f, g, and h stand for continuous maps. Again, in the category of all groups, X, Y, and Z stand for groups, f, g, and h for homomorphisms.

Less formally, composing maps is like following directed paths from one object to another (e.g., from set to set). In general, a diagram is commutative iff any two paths along morphisms that start at the same point and finish at the same point yield the same 'homomorphism' via compositions along successive morphisms. Commutativity of the whole diagram follows from commutativity

---

[17] The group product $f * g$ of loop f and loop g is given by the path of f followed by the path of g. The identity element is represented by the constant path, and the inverse f^{-1} of f is given by traversing f in the opposite direction. The fundamental group $\pi_1(X)$ is independent of the choice of base point x_0 because any loop through x_0 is homotopic to a loop through any other point x_1.

of its triangular components (depicting a ‘commutative flow’, see Figure 1). Study of commutative diagrams is popularly called ‘diagram chasing’, and provides a powerful tool for mathematical thought.

Many properties of mathematical constructions may be represented by *universal properties* of diagrams [52]. Consider the *Cartesian product* $X \times Y$ of two sets, consisting as usual of all ordered pairs $\langle x, y\rangle$ of elements $x \in X$ and $y \in Y$. The projections $\langle x, y\rangle \mapsto x$, $\langle x, y\rangle \mapsto y$ of the product on its ‘axes’ X and Y are functions $p : X \times Y \to X$, $q : X \times Y \to Y$. Any function $h : W \to X \times Y$ from a third set W is uniquely determined by its composites $p \circ h$ and $q \circ h$. Conversely, given W and two functions f and g as in the diagram below, there is a unique function h which makes the following diagram commute:

$$\begin{tikzcd}
& W \arrow[dl, "f"'] \arrow[d, "h"] \arrow[dr, "g"] & \\
X & X \times Y \arrow[l, "p"] \arrow[r, "q"'] & Y
\end{tikzcd}$$

This property describes the Cartesian product $X \times Y$ uniquely; the same diagram, read in the category of topological spaces or of groups, describes uniquely the Cartesian product of spaces or of the direct product of groups.

The construction ‘Cartesian product’ is technically called a ‘functor’ because it applies suitably both to the sets and to the functions between them; two functions $k : X \to X'$ and $l : Y \to Y'$ have a function $k \times l$ as their Cartesian product:

$$k \times l : X \times Y \to X' \times Y', \qquad \langle x, y\rangle \mapsto \langle kx, ly\rangle.$$

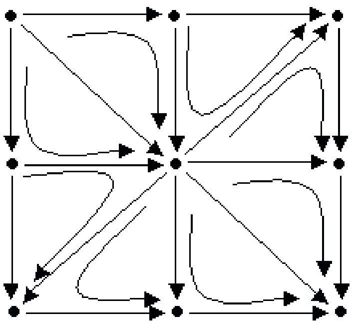


Fig. 1. A commutative flow (denoted by curved morphisms) on a triangulated digraph. Commutativity of the whole diagram follows from commutativity of its triangular components.

2.4 Categories

A category is a generic mathematical structure[18] consisting of a collection of *objects* (sets with possibly additional structure), with a corresponding collection of *morphisms*, or *morphisms*, between objects (agreeing with this additional structure). A category $\mathcal{K}$ is defined as a pair $(\mathtt{Ob}(\mathcal{K}), \mathtt{Mor}(\mathcal{K}))$ of generic objects $A, B, \ldots$ in $\mathtt{Ob}(\mathcal{K})$ and generic morphisms $f : A \to B$, $g : B \to C, \ldots$ in $\mathtt{Mor}(\mathcal{K})$ between objects, with *associative composition*:

$$A \xrightarrow{f} B \xrightarrow{g} C = A \xrightarrow{g \circ f} C,$$

and *identity* (*loop*) morphism. (Note that in topological literature, $\mathtt{Hom}(\mathcal{K})$ or $\mathtt{hom}(\mathcal{K})$ is used instead of $\mathtt{Mor}(\mathcal{K})$; see [59]).

A category $\mathcal{K}$ is usually depicted as a *commutative diagram* (i.e., a diagram with a common *initial object* A and *final object* D):

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "h"'] \arrow[dr, phantom, "\mathcal{K}"] & B \arrow[d, "g"] \\
C \arrow[r, "k"'] & D
\end{tikzcd}$$

To make this more precise, we say that a *category* $\mathcal{K}$ is defined if we have:

1. A *class of objects* $\{A, B, C, ...\}$ of $\mathcal{K}$, denoted by $\mathtt{Ob}(\mathcal{K})$;
2. A *set of morphisms*, or *morphisms* $\mathtt{Mor}_{\mathcal{K}}(A, B)$, with elements $f : A \to B$, defined for any *ordered pair* $(A, B) \in \mathcal{K}$, such that for two different pairs $(A, B) \neq (C, D)$ in $\mathcal{K}$, we have $\mathtt{Mor}_{\mathcal{K}}(A, B) \cap \mathtt{Mor}_{\mathcal{K}}(C, D) = \emptyset$;
3. For any *triplet* $(A, B, C) \in \mathcal{K}$ with $f : A \to B$ and $g : B \to C$, there is a *composition* of morphisms

$$\mathtt{Mor}_{\mathcal{K}}(B, C) \times \mathtt{Mor}_{\mathcal{K}}(A, B) \ni (g, f) \to g \circ f \in \mathtt{Mor}_{\mathcal{K}}(A, C),$$

written schematically as

$$\frac{f : A \to B, \qquad g : B \to C}{g \circ f : A \to C}.$$

Recall from above that if we have a morphism $f \in \mathtt{Mor}_{\mathcal{K}}(A, B)$, (otherwise written $f : A \to B$, or $A \xrightarrow{f} B$), then $A = \mathrm{dom}(f)$ is a *domain* of f, and $B = \mathrm{cod}(f)$ is a *codomain* of f (of which *range* of f is a subset, $B = \mathrm{ran}(f)$).

---

[18] We emphasize here that category theory, the abstract of the abstract, finds very important practical significance in terms of *type theory,* which is fundamental in building (among other things) programming languages that prevent bad things from being allowed while also enabling the expression of abstract concepts. Type checkers of languages are actually theorem provers!

To make $\mathcal{K}$ a category, it must also fulfill the following two properties:

1. *Associativity of morphisms*: for all $f \in \mathtt{Mor}_{\mathcal{K}}(A,B)$, $g \in \mathtt{Mor}_{\mathcal{K}}(B,C)$, and $h \in \mathtt{Mor}_{\mathcal{K}}(C,D)$, we have $h \circ (g \circ f) = (h \circ g) \circ f$; in other words, the following diagram is commutative

$$\begin{tikzcd}
A \arrow[r, "h \circ (g \circ f) = (h \circ g) \circ f"] \arrow[d, "f"'] & D \\
B \arrow[r, "g"'] & C \arrow[u, "h"']
\end{tikzcd}$$

2. *Existence of identity morphism*: for every object $A \in \mathtt{Ob}(\mathcal{K})$ exists a unique identity morphism $1_A \in \mathtt{Mor}_{\mathcal{K}}(A,A)$; for any two morphisms $f \in \mathtt{Mor}_{\mathcal{K}}(A,B)$, and $g \in \mathtt{Mor}_{\mathcal{K}}(B,C)$, compositions with identity morphism $1_B \in \mathtt{Mor}_{\mathcal{K}}(B,B)$ give $1_B \circ f = f$ and $g \circ 1_B = g$, i.e., the following diagram is commutative:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[dr, "f"'] & B \arrow[r, "g"] \arrow[d, "1_B"] & C \\
& B \arrow[ur, "g"'] &
\end{tikzcd}$$

The set of all morphisms of the category $\mathcal{K}$ is denoted

$$\mathtt{Mor}(\mathcal{K}) = \bigcup_{A,B \in Ob(\mathcal{K})} \mathtt{Mor}_{\mathcal{K}}(A,B).$$

If for two morphisms $f \in \mathtt{Mor}_{\mathcal{K}}(A,B)$ and $g \in \mathtt{Mor}_{\mathcal{K}}(B,A)$ the equality $g \circ f = 1_A$ is valid, then the morphism g is said to be *left inverse* (or *retraction*), of f, and f *right inverse* (or *section*) of g. A morphism which is both right and left inverse of f is said to be *two–sided inverse* of f.

A morphism $m : A \to B$ is called *monomorphism* in $\mathcal{K}$ (i.e., *1–1*, or *injection* map), if for any two parallel morphisms $f_1, f_2 : C \to A$ in $\mathcal{K}$ the equality $m \circ f_1 = m \circ f_2$ implies $f_1 = f_2$; in other words, m is monomorphism if it is *left cancellable.* Any morphism with a left inverse is monomorphism.

A morphism $e : A \to B$ is called *epimorphism* in $\mathcal{K}$ (i.e., *onto*, or *surjection* map), if for any two morphisms $g_1, g_2 : B \to C$ in $\mathcal{K}$ the equality $g_1 \circ e = g_2 \circ e$ implies $g_1 = g_2$; in other words, e is epimorphism if it is *right cancellable.* Any morphism with a right inverse is epimorphism.

A morphism $f : A \to B$ is called *isomorphism* in $\mathcal{K}$ (denoted as $f : A \cong B$) if there exists a morphism $f^{-1} : B \to A$ which is a two–sided inverse of f in $\mathcal{K}$. The relation of isomorphism is reflexive, symmetric, and transitive, that is, an equivalence relation.

For example, an isomorphism in the category of sets is called a set–isomorphism, or a *bijection*, in the category of topological spaces is called a topological isomorphism, or a *homeomorphism*, in the category of differentiable manifolds is called a differentiable isomorphism, or a *diffeomorphism.*

A morphism $f \in \mathtt{Mor}_{\mathcal{K}}(A, B)$ is *regular* if there exists a morphism $g : B \to A$ in $\mathcal{K}$ such that $f \circ g \circ f = f$. Any morphism with either a left or a right inverse is regular.

An object T is a *terminal object* in $\mathcal{K}$ if to each object $A \in \mathtt{Ob}(\mathcal{K})$ there is exactly one morphism $A \to T$. An object S is an *initial object* in $\mathcal{K}$ if to each object $A \in \mathtt{Ob}(\mathcal{K})$ there is exactly one morphism $S \to A$. A *null object* $Z \in \mathtt{Ob}(\mathcal{K})$ is an object which is both initial and terminal; it is unique up to isomorphism. For any two objects $A, B \in \mathtt{Ob}(\mathcal{K})$ there is a unique morphism $A \to Z \to B$ (the composite through Z), called the *zero morphism* from A to B.

A notion of subcategory is analogous to the notion of subset. A subcategory $\mathcal{L}$ of a category $\mathcal{K}$ is said to be a *complete subcategory* iff for any objects $A, B \in \mathcal{L}$, every morphism $A \to B$ of $\mathcal{L}$ is in $\mathcal{K}$.

A *groupoid* is a category in which every morphism is invertible. A typical groupoid is the *fundamental groupoid* $\Pi_1(X)$ of a topological space X. An object of $\Pi_1(X)$ is a point $x \in X$, and a morphism $x \to x'$ of $\Pi_1(X)$ is a homotopy class of paths f from x to x'. The *composition* of paths $g : x' \to x''$ and $f : x \to x'$ is the path h which is 'f followed by g'. Composition applies also to homotopy classes, and makes $\Pi_1(X)$ a category and a groupoid (the inverse of any path is the same path traced in the opposite direction).

A *group* is a groupoid with one object, i.e., a *category with one object* in which *all morphisms are isomorphisms.* Therefore, if we try to generalize the concept of a group, keeping associativity as an essential property, we get the notion of a category.

A category is *discrete* if every morphism is an identity. A *monoid* is a category with one object, which is a group without inverses. A group is a category with one object in which every morphism has a two–sided inverse under composition.

Homological algebra was the progenitor of category theory (see e.g., [62]). Generalizing L. Euler's formula: $f + v = e + 2$, for the faces f, vertices v and edges e of a convex polyhedron, E. Betti defined *numerical invariants of spaces* by formal addition and subtraction of faces of various dimensions. H. Poincaré formalized these and introduced the concept of *homology.* E. Noether stressed the fact that these calculations go on in Abelian groups, and that the operation ∂_n taking a face of dimension n to the alternating sum of faces of dimension $n-1$ which form its boundary is a homomorphism, and it also satisfies *the boundary of a boundary is zero* rule: $\partial_n \circ \partial_{n+1} = 0$. There are many ways of approximating a given space by polyhedra, but the quotient $H_n = \operatorname{Ker} \partial_n / \operatorname{Im} \partial_{n+1}$ is an invariant, the *homology group.*

As a physical example from [64, 65], consider some physical system of type A (e.g., an electron) and perform some *physical operation* f on it (e.g., perform a measurement on it), which results in a possibly different system B (e.g., a perturbed electron), thus having a map $f : A \longrightarrow B$. In a same way, we can perform a consecutive operation $g : B \longrightarrow C$ (e.g., perform the second measurement, this time on B), possibly resulting in a different system C (e.g., a secondly perturbed electron). Thus, we have a composition: $k = g \circ f$, representing the consecutive application of these two physical operations, or the following diagram commutes:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[dr, "k"'] & B \arrow[d, "g"] \\
 & C
\end{tikzcd}$$

In a similar way, we can perform another consecutive operation $h : C \longrightarrow D$ (e.g., perform the third measurement, this time on C), possibly resulting in a different system D (e.g., a thirdly perturbed electron). Clearly we have an associative composition $(h \circ g) \circ f = h \circ (g \circ f)$, or the following diagram commutes:

$$\begin{tikzcd}
A \arrow[r, "h \circ (g \circ f) = (h \circ g) \circ f"] \arrow[d, "f"'] & D \\
B \arrow[r, "g"'] & C \arrow[u, "h"']
\end{tikzcd}$$

Finally, if we introduce a trivial operation $1_A \in \mathtt{Mor}_{\mathcal{K}}(A, A)$, meaning 'doing nothing on a system of type A', we have $1_B \circ f = f \circ 1_A = f$. In this way, we have constructed a generic physical category (for more details, see [64, 65]).

For the same *operational* reasons, categories could be expected to play an important role in other fields where operations/processes play a central role: e.g., Computer Science (computer programs as morphisms) and Logic & Proof Theory (proofs as morphisms). In the theoretical counterparts to these fields category theory has become quite common practice (see [66]).

2.5 Functors

In algebraic topology, one attempts to assign to every topological space X some algebraic object $\mathcal{F}(X)$ in such a way that to every C^0-function $f : X \to Y$ there is assigned a homomorphism $\mathcal{F}(f) : \mathcal{F}(X) \longrightarrow \mathcal{F}(Y)$ (see [59, 3]). One advantage of this procedure is, e.g., that if one is trying to prove the non–existence of a C^0-function $f : X \to Y$ with certain properties, one may find it relatively easy to prove the non–existence of the corresponding algebraic function $\mathcal{F}(f)$ and hence deduce that f could not exist. In other words, $\mathcal{F}$ is to be a 'homomorphism' from one category (e.g., $\mathcal{T}$) to another (e.g., $\mathcal{G}$ or $\mathcal{A}$). Formalization of this notion is a *functor*.

A functor is a generic *picture* projecting (all objects and morphisms of) a source category into a target category. Let $\mathcal{K} = (\mathtt{Ob}(\mathcal{K}), \mathtt{Mor}(\mathcal{K}))$ be a *source* (or domain) *category* and $\mathcal{L} = (\mathtt{Ob}(\mathcal{L}), \mathtt{Mor}(\mathcal{L}))$ be a *target* (or codomain) category. A functor $\mathcal{F} = (\mathcal{F}_O, \mathcal{F}_M)$ is defined as a pair of maps, $\mathcal{F}_O : \mathtt{Ob}(\mathcal{K}) \to \mathtt{Ob}(\mathcal{L})$ and $\mathcal{F}_M : \mathtt{Mor}(\mathcal{K}) \to \mathtt{Mor}(\mathcal{L})$, preserving categorical symmetry (i.e., commutativity of all diagrams) of $\mathcal{K}$ in $\mathcal{L}$.

More precisely, a *covariant functor*, or simply a *functor*, $\mathcal{F}_* : \mathcal{K} \to \mathcal{L}$ is a *picture* in the target category $\mathcal{L}$ of (all objects and morphisms of) the source category $\mathcal{K}$:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "h"'] \arrow[dr, phantom, "\mathcal{K}"] & B \arrow[d, "g"] \\
C \arrow[r, "k"'] & D
\end{tikzcd}
\quad \xrightarrow{\mathcal{F}_*} \quad
\begin{tikzcd}
\mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "\mathcal{F}(h)"'] \arrow[dr, phantom, "\mathcal{L}"] & \mathcal{F}(B) \arrow[d, "\mathcal{F}(g)"] \\
\mathcal{F}(C) \arrow[r, "\mathcal{F}(k)"'] & \mathcal{F}(D)
\end{tikzcd}$$

Similarly, a *contravariant functor*, or a *cofunctor*, $\mathcal{F}^* : \mathcal{K} \to \mathcal{L}$ is a *dual picture* with reversed morphisms:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "h"'] \arrow[dr, phantom, "\mathcal{K}"] & B \arrow[d, "g"] \\
C \arrow[r, "k"'] & D
\end{tikzcd}
\quad \xrightarrow{\mathcal{F}^*} \quad
\begin{tikzcd}
\mathcal{F}(A) \arrow[dr, phantom, "\mathcal{L}"] & \mathcal{F}(B) \arrow[l, "\mathcal{F}(f)"'] \\
\mathcal{F}(C) \arrow[u, "\mathcal{F}(h)"] & \mathcal{F}(D) \arrow[l, "\mathcal{F}(k)"] \arrow[u, "\mathcal{F}(g)"']
\end{tikzcd}$$

In other words, a *functor* $\mathcal{F} : \mathcal{K} \to \mathcal{L}$ from a *source* category $\mathcal{K}$ to a *target* category $\mathcal{L}$, is a pair $\mathcal{F} = (\mathcal{F}_O, \mathcal{F}_M)$ of maps $\mathcal{F}_O : \mathtt{Ob}(\mathcal{K}) \to \mathtt{Ob}(\mathcal{L})$, $\mathcal{F}_M : \mathtt{Mor}(\mathcal{K}) \to \mathtt{Mor}(\mathcal{L})$, such that

1. If $f \in \mathtt{Mor}_{\mathcal{K}}(A, B)$ then $\mathcal{F}_M(f) \in \mathtt{Mor}_{\mathcal{L}}(\mathcal{F}_O(A), \mathcal{F}_O(B))$ in case of the *covariant* functor $\mathcal{F}_*$, and $\mathcal{F}_M(f) \in \mathtt{Mor}_{\mathcal{L}}(\mathcal{F}_O(B), \mathcal{F}_O(A))$ in case of the *contravariant* functor $\mathcal{F}^*$;
2. For all $A \in \mathtt{Ob}(\mathcal{K}) : \mathcal{F}_M(1_A) = 1_{\mathcal{F}_O(A)}$;
3. For all $f, g \in \mathtt{Mor}(\mathcal{K})$: if $\mathrm{cod}(f) = \mathrm{dom}(g)$, then $\mathcal{F}_M(g \circ f) = \mathcal{F}_M(g) \circ \mathcal{F}_M(f)$ in case of the *covariant* functor $\mathcal{F}_*$, and $\mathcal{F}_M(g \circ f) = \mathcal{F}_M(f) \circ \mathcal{F}_M(g)$ in case of the *contravariant* functor $\mathcal{F}^*$.

Category theory originated in algebraic topology, which tried to assign algebraic invariants to topological structures. The golden rule of such *invariants* is that they should be *functors*.

For example, in computer science, every monad is a functor, but not every functor is a monad. Monads are vitally important in creating higher-order functions that capture notions of computation such as sequence and control of flow. So, we can think of a functor as a structure that can be mapped over using a mapping that takes on its ‘elements’ to give a new structure.

As a standard topological example, the *fundamental group* π_1 is a functor. Algebraic topology constructs a group called the *fundamental group* $\pi_1(X)$

from any topological space X, which keeps track of how many holes the space X has. But also, any map between topological spaces determines a homomorphism $\phi : \pi_1(X) \to \pi_1(Y)$ of the fundamental groups. So the fundamental group is really a functor $\pi_1 : \mathcal{T} \to \mathcal{G}$. This allows us to completely transpose any situation involving *spaces* and *continuous maps* between them to a parallel situation involving *groups* and *homomorphisms* between them, and thus reduce some topology problems to algebra problems.

Also, singular homology in a given dimension n assigns to each topological space X an Abelian group $H_n(X)$, its nth *homology group* of X, and also to each continuous map $f : X \to Y$ of spaces a corresponding homomorphism $H_n(f) : H_n(X) \to H_n(Y)$ of groups, and this in such a way that $H_n(X)$ becomes a functor $H_n : \mathcal{T} \to \mathcal{A}$.

The leading idea in the *use of functors in topology* is that H_n or π_n gives an algebraic picture or image not just of the topological spaces X, Y but also of all the continuous maps $f : X \to Y$ between them.

Similarly, there is a functor $\varPi_1 : \mathcal{T} \to \mathcal{G}$, called the 'fundamental groupoid functor', which plays a very basic role in algebraic topology. Here's how we get from any space X its 'fundamental groupoid' $\varPi_1(X)$. To say what the groupoid $\varPi_1(X)$ is, we need to say what its objects and morphisms are. The objects in $\varPi_1(X)$ are just the *points* of X and the morphisms are just certain equivalence classes of *paths* in X. More precisely, a morphism $f : x \to y$ in $\varPi_1(X)$ is just an equivalence class of continuous paths from x to y, where two paths from x to y are decreed equivalent if one can be continuously deformed to the other while not moving the endpoints. (If this equivalence relation holds, we say the two paths are 'homotopic', and we call the equivalence classes 'homotopy classes of paths' [59].

Another examples are covariant *forgetful* functors:

- From the category of topological spaces to the category of sets; it 'forgets' the topology–structure.
- From the category of metric spaces to the category of topological spaces with the topology induced by the metrics; it 'forgets' the metric.

For each category $\mathcal{K}$, the *identity functor* $I_{\mathcal{K}}$ takes every $\mathcal{K}-$object and every $\mathcal{K}-$morphism to itself.

Given a category $\mathcal{K}$ and its subcategory $\mathcal{L}$, we have an *inclusion functor* $\mathtt{In} : \mathcal{L} \longrightarrow \mathcal{K}$.

Given a category $\mathcal{K}$, a *diagonal functor* $\varDelta : \mathcal{K} \longrightarrow \mathcal{K} \times \mathcal{K}$ takes each object $A \in \mathcal{K}$ to the object (A, A) in the product category $\mathcal{K} \times \mathcal{K}$.

Given a category $\mathcal{K}$ and a category of sets $\mathcal{S}$, each object $A \in \mathcal{K}$ determines a *covariant Hom–functor* $\mathcal{K}[A, \_] : \mathcal{K} \to \mathcal{S}$, a *contravariant Hom–functor* $\mathcal{K}[\_, A] : \mathcal{K} \longrightarrow \mathcal{S}$, and a *Hom–bifunctor* $\mathcal{K}[\_, \_] : \mathcal{K}^{op} \times \mathcal{K} \to \mathcal{S}$.

A functor $\mathcal{F} : \mathcal{K} \to \mathcal{L}$ is a *faithful functor* if for all $A, B \in \mathtt{Ob}(\mathcal{K})$ and for all $f, g \in \mathtt{Mor}_{\mathcal{K}}(A, B)$, $\mathcal{F}(f) = \mathcal{F}(g)$ implies $f = g$; it is a *full functor* if for every $h \in \mathtt{Mor}_{\mathcal{L}}(\mathcal{F}(A), \mathcal{F}(B))$, there is $g \in \mathtt{Mor}_{\mathcal{K}}(A, B)$ such that $h = \mathcal{F}(g)$; it is a *full embedding* if it is both full and faithful.

A *representation of a group* is a functor $\mathcal{F} : \mathcal{G} \to \mathcal{V}$. Thus, a category is a generalization of a group and group representations are a special case of category representations.

2.6 Natural Transformations

A *natural transformation* (i.e., a *functor morphism*) $\boldsymbol{\tau} : \mathcal{F} \stackrel{\cdot}{\to} \mathcal{G}$ is a *map between two functors of the same variance*, $(\mathcal{F}, \mathcal{G}) : \mathcal{K} \rightrightarrows \mathcal{L}$, preserving categorical symmetry:

$$\begin{tikzcd}
A \arrow[r, "f"] & B & {} \arrow[r, "\mathcal{F}"] \arrow[d, phantom, "\boldsymbol{\tau}\Downarrow"] & {} & \mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "\boldsymbol{\tau}_A"'] \arrow[dr, phantom, "\mathcal{L}"] & \mathcal{F}(B) \arrow[d, "\boldsymbol{\tau}_B"] \\
\mathcal{K} & & {} \arrow[r, "\mathcal{G}"] & {} & \mathcal{G}(A) \arrow[r, "\mathcal{G}(f)"'] & \mathcal{G}(B)
\end{tikzcd}$$

More precisely, all functors of the same variance from a source category $\mathcal{K}$ to a target category $\mathcal{L}$ form themselves objects of the *functor category* $\mathcal{L}^{\mathcal{K}}$. Morphisms of $\mathcal{L}^{\mathcal{K}}$, called *natural transformations*, are defined as follows.

Let $\mathcal{F} : \mathcal{K} \to \mathcal{L}$ and $\mathcal{G} : \mathcal{K} \to \mathcal{L}$ be two functors of the same variance from a category $\mathcal{K}$ to a category $\mathcal{L}$. Natural transformation $\mathcal{F} \xrightarrow{\boldsymbol{\tau}} \mathcal{G}$ is a family of morphisms such that for all $f \in \mathtt{Mor}_{\mathcal{K}}(A, B)$ in the source category $\mathcal{K}$, we have $\mathcal{G}(f) \circ \boldsymbol{\tau}_A = \boldsymbol{\tau}_B \circ \mathcal{F}(f)$ in the target category $\mathcal{L}$. Then we say that the *component* $\boldsymbol{\tau}_A : \mathcal{F}(A) \to \mathcal{G}(A)$ *is natural in* A.

If we think of a functor $\mathcal{F}$ as giving a *picture* in the target category $\mathcal{L}$ of (all the objects and morphisms of) the source category $\mathcal{K}$, then a natural transformation $\boldsymbol{\tau}$ represents a set of morphisms mapping the picture $\mathcal{F}$ to another picture $\mathcal{G}$, preserving the commutativity of all diagrams.

An invertible natural transformation, such that all components $\boldsymbol{\tau}_A$ are isomorphisms) is called a *natural equivalence* (or, *natural isomorphism*). In this case, the inverses $(\boldsymbol{\tau}_A)^{-1}$ in $\mathcal{L}$ are the components of a natural isomorphism $(\boldsymbol{\tau})^{-1} : \mathcal{G} \xrightarrow{*} \mathcal{F}$. Natural equivalences are among the most important *metamathematical constructions* in algebraic topology (see [59]).

As a mathematical example, let $\mathcal{B}$ be the category of Banach spaces over $\mathbb{R}$ and bounded linear maps. Define $D : \mathcal{B} \to \mathcal{B}$ by taking $D(X) = X^* =$ Banach space of bounded linear functionals on a space X and $D(f) = f^*$ for $f : X \to Y$ a bounded linear map. Then D is a cofunctor. $D^2 = D \circ D$ is also a functor. We also have the identity functor $1 : \mathcal{B} \to \mathcal{B}$. Define $T : 1 \to D \circ D$ as follows: for every $X \in \mathcal{B}$ let $T(X) : X \to D^2X = X^{**}$ be the *natural inclusion* – that is, for $x \in X$ we have $[T(X)(x)](f) = f(x)$ for every $f \in X^*$. T is a natural transformation. On the subcategory of nD Banach spaces T is even a natural equivalence. The largest subcategory of $\mathcal{B}$ on which T is a natural equivalence is called the category of reflexive Banach spaces [59].

As a physical example, when we want to be able to conceive two physical systems A and B as one whole (see [64, 65]), we can denote this using a (symmetric) monoidal tensor product $A \otimes B$, and hence also need to consider the compound operations

$$A \otimes B \xrightarrow{f \otimes g} C \otimes D,$$

inherited from the operations on the individual systems. Now, a (symmetric) *monoidal category* is a category $\mathcal{K}$ defined as a pair $(\mathtt{Ob}(\mathcal{K}), \mathtt{Mor}(\mathcal{K}))$ of generic objects $A, B, \ldots$ in $\mathtt{Ob}(\mathcal{K})$ and generic morphisms $f : A \to B$, $g : B \to C, \ldots$ in $\mathtt{Mor}(\mathcal{K})$ between objects, defined using the symmetric monoidal tensor product:

$$\begin{aligned} \mathtt{Ob}(\mathcal{K}) : \{A, B\} &\mapsto A \otimes B, \\ \mathtt{Mor}(\mathcal{K}) : \{A \xrightarrow{\quad f \quad} B, C \xrightarrow{\quad g \quad} D\} &\mapsto A \otimes C \xrightarrow{f \otimes g} B \otimes D, \end{aligned}$$

with the additional notion of *bifunctoriality*: if we apply an operation f to one system and an operation g to another system, then the order in which we apply them does not matter; that is, the following diagram commutes:

$$\begin{tikzcd} A_1 \otimes A_2 \arrow[r, "f \otimes 1_{A_2}"] \arrow[d, "1_{A_1} \otimes g"'] & B_1 \otimes A_2 \\ A_1 \otimes B_2 \arrow[r, "f \otimes 1_{B_2}"'] & B_1 \otimes B_2 \arrow[u, "1_{B_1} \otimes g"'] \end{tikzcd}$$

which shows that both paths yield the same result (see [64, 65] for technical details).

As 'categorical fathers', S. Eilenberg and S. MacLane, first observed, 'category' has been defined in order to define 'functor' and 'functor' has been defined in order to define 'natural transformations' [58, 52]).

Natural transformations can be *composed* in two different ways. First, we have an 'ordinary' composition: if $\mathcal{F}, \mathcal{G}$ and $\mathcal{H}$ are three functors from the source category $\mathcal{A}$ to the target category $\mathcal{B}$, and then $\boldsymbol{\alpha} : \mathcal{F} \stackrel{\cdot}{\to} \mathcal{G}$, $\boldsymbol{\beta} : \mathcal{G} \stackrel{\cdot}{\to} \mathcal{H}$ are two natural transformations, then the formula

$$(\beta \circ \alpha)_A = \beta_A \circ \alpha_A, \qquad \text{(for all } A \in \mathcal{A}), \tag{2}$$

defines a new natural transformation $\beta \circ \alpha : \mathcal{F} \stackrel{\cdot}{\to} \mathcal{H}$. This composition law is clearly associative and possesses a unit $1_{\mathcal{F}}$ at each functor $\mathcal{F}$, whose $\mathcal{A}$–component is $1_{\mathcal{F}\mathcal{A}}$.

Second, we have the *Godement product* of natural transformations, usually denoted by $*$. Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be three categories, $\mathcal{F}, \mathcal{G}$, $\mathcal{H}$ and $\mathcal{K}$ be four functors such that $(\mathcal{F}, \mathcal{G}) : \mathcal{A} \rightrightarrows \mathcal{B}$ and $(\mathcal{H}, \mathcal{K}) : \mathcal{B} \rightrightarrows \mathcal{C}$, and $\boldsymbol{\alpha} : \mathcal{F} \stackrel{\cdot}{\to} \mathcal{G}$, $\boldsymbol{\beta} : \mathcal{H} \stackrel{\cdot}{\to} \mathcal{K}$ be two natural transformations. Now, instead of (2), the Godement composition is given by

$$(\beta * \alpha)_A = \beta_{GA} \circ H(\alpha_A) = K(\alpha_A) \circ \beta_{FA}, \qquad \text{(for all } A \in \mathcal{A}), \tag{3}$$

which defines a new natural transformation $\beta * \alpha : \mathcal{H} \circ \mathcal{F} \stackrel{\cdot}{\to} \mathcal{K} \circ \mathcal{G}$.

Finally, the two compositions (2) and (3) of natural transformations can be combined as

$$(\delta * \gamma) \circ (\beta * \alpha) = (\delta \circ \beta) * (\gamma \circ \alpha)\,,$$

where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are three categories, $\mathcal{F}, \mathcal{G}$, $\mathcal{H}$, $\mathcal{K}$, $\mathcal{L}$, $\mathcal{M}$ are six functors, and $\boldsymbol{\alpha} : \mathcal{F} \overset{\cdot}{\to} \mathcal{H}$, $\boldsymbol{\beta} : \mathcal{G} \overset{\cdot}{\to} \mathcal{K}$, $\boldsymbol{\gamma} : \mathcal{H} \overset{\cdot}{\to} \mathcal{L}$, $\boldsymbol{\delta} : \mathcal{K} \overset{\cdot}{\to} \mathcal{M}$ are four natural transformations.[19]

---

[19] Double natural transformations are called *dinatural transformations*. An *end of a functor* $S : C^{op} \times C \to X$ is a universal dinatural transformation from a constant e to S. In other words, an end of S is a pair $\langle e, \omega \rangle$, where e is an object of X and $\omega : e \overset{\cdot\cdot}{\to} S$ is a *wedge (dinatural) transformation* with the property that to every wedge $\beta : x \overset{\cdot\cdot}{\to} S$ there is a unique morphism $h : x \to e$ of B with $\beta_c = \omega_c h$ for all $a \in C$. We call ω the *ending wedge* with *components* ω_c, while the object e itself, by abuse of language, is called the end of S and written with integral notation as $\int_c S(c,c)$; thus

$$S(c,c) \xrightarrow{\omega_c} \int_c S(c,c) = e.$$

Note that the 'variable of integration' c appears twice under the integral sign (once contravariant, once covariant) and is 'bound' by the integral sign, in that the result no longer depends on c and so is unchanged if 'c' is replaced by any other letter standing for an object of the category C. These properties are like those of the letter x under the usual integral symbol $\int f(x)\,dx$ of calculus.

Every end is manifestly a limit (see below) – specifically, a limit of a suitable diagram in X made up of pieces like $S(b,b) \to S(b,c) \to S(c,c)$.

For each functor $T : C \to X$ there is an isomorphism

$$\int_c S(c,c) = \int_c Tc \cong \operatorname{Lim} T,$$

valid when either the end of the limit exists, carrying the ending wedge to the limiting cone; the indicated notation thus allows us to write any limit as an integral (an end) without explicitly mentioning the dummy variable (the first variable c of S).

A functor $H : X \to Y$ is said to *preserve the end* of a functor $S : C^{op} \times C \to X$ when $\omega : e \overset{\cdot\cdot}{\to} S$ an end of S in X implies that $H\omega : He \overset{\cdot\cdot}{\to} HS$ is an and for HS; in symbols

$$H \int_c S(c,c) = \int_c HS(c,c).$$

Similarly, H *creates* the end of S when to each end $v : y \overset{\cdot\cdot}{\to} HS$ in Y there is a unique wedge $\omega : e \overset{\cdot\cdot}{\to} S$ with $H\omega = v$, and this wedge ω is an end of S.

The definition of the coend of a functor $S : C^{op} \times C \to X$ is dual to that of an end. A *coend* of S is a pair $\langle d, \zeta \rangle$, consisting of an object $d \in X$ and a wedge $\zeta : S \overset{\cdot\cdot}{\to} d$. The object d (when it exists, unique up to isomorphism) will usually be written with an integral sign and with the bound variable c as superscript; thus

$$S(c,c) \xrightarrow{\zeta_c} \int^c S(c,c) = d.$$

The formal properties of coends are dual to those of ends. Both are much like those for integrals in calculus (see [52], for technical details).

2.7 Limits and Colimits

In abstract algebra constructions are often defined by an abstract property which requires the existence of unique morphisms under certain conditions. These properties are called *universal properties.* The *limit* of a functor generalizes the notions of inverse limit and product used in various parts of mathematics. The dual notion, *colimit*, generalizes direct limits and direct sums. Limits and colimits are defined via universal properties and provide many examples of *adjoint functors.*

A *limit* of a covariant functor $\mathcal{F} : \mathcal{J} \to \mathcal{C}$ is an object L of $\mathcal{C}$, together with morphisms $\phi_X : L \to \mathcal{F}(X)$ for every object X of $\mathcal{J}$, such that for every morphism $f : X \to Y$ in $\mathcal{J}$, we have $\mathcal{F}(f)\phi_X = \phi_Y$, and such that the following *universal property* is satisfied: for any object N of $\mathcal{C}$ and any set of morphisms $\psi_X : N \to \mathcal{F}(X)$ such that for every morphism $f : X \to Y$ in $\mathcal{J}$, we have $\mathcal{F}(f)\psi_X = \psi_Y$, there exists precisely one morphism $u : N \to L$ such that $\phi_X u = \psi_X$ for all X. If $\mathcal{F}$ has a limit (which it need not), then the limit is defined up to a unique isomorphism, and is denoted by $\lim \mathcal{F}$.

Analogously, a *colimit* of the functor $\mathcal{F} : \mathcal{J} \to \mathcal{C}$ is an object L of $\mathcal{C}$, together with morphisms $\phi_X : \mathcal{F}(X) \to L$ for every object X of $\mathcal{J}$, such that for every morphism $f : X \to Y$ in $\mathcal{J}$, we have $\phi_Y \mathcal{F}(X) = \phi_X$, and such that the following universal property is satisfied: for any object N of $\mathcal{C}$ and any set of morphisms $\psi_X : \mathcal{F}(X) \to N$ such that for every morphism $f : X \to Y$ in $\mathcal{J}$, we have $\psi_Y \mathcal{F}(X) = \psi_X$, there exists precisely one morphism $u : L \to N$ such that $u\phi_X = \psi_X$ for all X. The colimit of $\mathcal{F}$, unique up to unique isomorphism if it exists, is denoted by $\operatorname{colim} \mathcal{F}$.

Limits and colimits are related as follows: A functor $\mathcal{F} : \mathcal{J} \to \mathcal{C}$ has a colimit iff for every object N of $\mathcal{C}$, the functor $X \longmapsto Mor_{\mathcal{C}}(\mathcal{F}(X), N)$ (which is a covariant functor on the dual category $\mathcal{J}^{op}$) has a limit. If that is the case, then $Mor_{\mathcal{C}}(\operatorname{colim} \mathcal{F}, N) = \lim Mor_{\mathcal{C}}(\mathcal{F}(-), N)$ for every object N of $\mathcal{C}$.

2.8 Adjunction

The most important functorial operation is adjunction; as S. MacLane once said, "Adjoint functors arise everywhere" [52].

The *adjunction* $\varphi : \mathcal{F} \dashv \mathcal{G}$ between two functors $(\mathcal{F}, \mathcal{G}) : \mathcal{K} \leftrightarrows \mathcal{L}$ of *opposite variance* [63], represents a *weak functorial inverse:*

$$\frac{f : \mathcal{F}(A) \to B}{\varphi(f) : A \to \mathcal{G}(B)},$$

forming a *natural equivalence* $\varphi : \mathtt{Mor}_{\mathcal{K}}(\mathcal{F}(A), B) \xrightarrow{\varphi} \mathtt{Mor}_{\mathcal{L}}(A, \mathcal{G}(B))$. The adjunction isomorphism is given by a *bijective correspondence* (a *1–1* and *onto* map on objects) $\varphi : \mathtt{Mor}(\mathcal{K}) \ni f \to \varphi(f) \in \mathtt{Mor}(\mathcal{L})$ of isomorphisms in the two categories, $\mathcal{K}$ (with a representative object A), and $\mathcal{L}$ (with a representative object B). It can be depicted as a *non–commutative diagram*

$$\begin{tikzcd}
\mathcal{F}(A) \arrow[dd, "f"'] & & & A \arrow[lll, "\mathcal{F}"'] \arrow[dd, "\varphi(f)"] \\
& \mathcal{K} & \mathcal{L} & \\
B \arrow[rrr, "\mathcal{G}"'] & & & \mathcal{G}(B)
\end{tikzcd}$$

In this case $\mathcal{F}$ is called *left adjoint*, while $\mathcal{G}$ is called *right adjoint.*

In other words, an *adjunction* $F \dashv G$ *between two functors* $(\mathcal{F}, \mathcal{G})$ of opposite variance, from a source category $\mathcal{K}$ to a target category $\mathcal{L}$, is denoted by $(\mathcal{F}, \mathcal{G}, \boldsymbol{\eta}, \boldsymbol{\varepsilon}) : \mathcal{K} \leftrightarrows \mathcal{L}$. Here, $\mathcal{F} : \mathcal{L} \to \mathcal{K}$ is the *left (upper) adjoint functor*, $\mathcal{G} : \mathcal{K} \leftarrow \mathcal{L}$ is the *right (lower) adjoint functor*, $\boldsymbol{\eta} : 1_{\mathcal{L}} \to \mathcal{G} \circ \mathcal{F}$ is the *unit natural transformation* (or, *front adjunction*), and $\boldsymbol{\varepsilon} : \mathcal{F} \circ \mathcal{G} \to 1_{\mathcal{K}}$ is the *counit natural transformation* (or, *back adjunction*).

For example, $\mathcal{K} = \mathcal{S}$ is the category of sets and $\mathcal{L} = \mathcal{G}$ is the category of groups. Then $\mathcal{F}$ turns any set into the *free group* on that set, while the 'forgetful' functor $\mathcal{F}^*$ turns any group into the *underlying set* of that group. Similarly, all sorts of other 'free' and 'underlying' constructions are also left and right adjoints, respectively.

Right adjoints preserve *limits*, and left adjoints preserve *colimits*.

The category $\mathcal{C}$ is called a *cocomplete category* if every functor $\mathcal{F} : \mathcal{J} \to \mathcal{C}$ has a colimit. The following categories are cocomplete: $\mathcal{S}, \mathcal{G}, \mathcal{A}, \mathcal{T}$, and $\mathcal{PT}$.

The importance of adjoint functors lies in the fact that every functor which has a left adjoint (and therefore is a right adjoint) is continuous. In the category $\mathcal{A}$ of Abelian groups, this shows e.g. that the kernel of a product of homomorphisms is naturally identified with the product of the kernels. Also, limit functors themselves are continuous. A covariant functor $\mathcal{F} : \mathcal{J} \to \mathcal{C}$ is *co-continuous* if it transforms colimits into colimits. Every functor which has a right adjoint (and therefore is a left adjoint) is co-continuous.

A Weak Physiological Example: Sensory–Motor Adjunction Sensations from the skin, muscles, and internal organs of the body, are transmitted to the central nervous system via axons that enter via spinal nerves. They are called *sensory pathways.* On the other hand, the motor system executes control over the skeletal muscles of the body via several major tracts (including pyramidal and extrapyramidal). They are called *motor pathways.* Sensory–motor (or, sensorimotor) control/coordination concerns relationships between sensation and movement or, more broadly, between perception and action. The interplay of sensory and motor processes provides the basis of observable human behavior. Anatomically, its top–level, association link can be visualized as a talk between sensory and motor Penfield's homunculi. This sensory–motor control system can be modelled as an adjunction between the afferent sensory functor $\mathcal{S} : \mathcal{BODY} \to \mathcal{BRAIN}$ and the efferent motor functor $\mathcal{M} : \mathcal{BRAIN} \to \mathcal{BODY}$. Thus, we have $\mathcal{SMC} : \mathcal{S} \dashv \mathcal{M}$, with $(\mathcal{S}, \mathcal{M}) : \mathcal{BRAIN} \leftrightarrows \mathcal{BODY}$ and depicted as

$$\begin{tikzcd}
\mathcal{S}(A) \arrow[d, "f"', "\text{BRAIN}"] & & A \arrow[ll, "\mathcal{S}"'] \arrow[d, "\text{BODY}"', "\mathcal{SMC}(f)"] \\
B \arrow[rr, "\mathcal{M}"'] & & \mathcal{M}(B)
\end{tikzcd}$$

This adjunction offers a mathematical answer to the fundamental question: How would *Nature* solve a general biodynamics control/coordination problem? *By using a weak functorial inverse of sensory neural pathways and motor neural pathways, Nature controls human behavior in general, and human motion in particular.*

More generally, normal functioning of human body is achieved through interplay of a number of physiological systems – Objects of the category BODY: musculoskeletal system, circulatory system, gastrointestinal system, integumentary system, urinary system, reproductive system, immune system and endocrine system. These systems are all interrelated, so one can say that the Morphisms between them make the proper functioning of the BODY as a whole. On the other hand, BRAIN contains the images of all above functional systems (Brain objects) and their interrelations (Brain morphisms), for the purpose of body control. This body–control performed by the brain is partly unconscious, through neuro–endocrine complex, and partly conscious, through neuro–muscular complex. A generalized sensory functor $\mathcal{SS}$ sends the information about the state of all Body objects (at any time instant) to their images in the Brain. A generalized motor functor $\mathcal{MM}$ responds to these upward sensory signals by sending downward corrective action–commands from the Brain's objects and morphisms to the Body's objects and morphisms. For physiological details, see [2]. For other bio–physical applications of categorical meta-language, see [1, 3, 4].

2.9 Groups and Related Algebraic Structures

A *group* is a pointed set (G, e) with a *multiplication* $\mu : G \times G \to G$ and an *inverse* $\nu : G \to G$ such that the following diagrams commute [59]:

1.

$$\begin{tikzcd}
G \arrow[r, "{(e,1)}"] \arrow[dr, "1"'] & G \times G \arrow[r, "{(1,e)}"] \arrow[d, "\mu"] & G \\
& G \arrow[ur, "1"'] &
\end{tikzcd}$$

(e is a two–sided identity)

2.

$$\begin{tikzcd}
G \times G \times G \arrow[r, "\mu \times 1"] \arrow[d, "1 \times \mu"'] & G \times G \arrow[d, "\mu"] \\
G \times G \arrow[r, "\mu"'] & G
\end{tikzcd}$$

(associativity)

3.

$$\begin{tikzcd}
G \arrow[r, "(\nu, 1)"] \arrow[dr, "e"'] & G \times G \arrow[r, "(1, \nu)"] \arrow[d, "\mu"] & G \\
 & G \arrow[ur, "e"'] &
\end{tikzcd}$$

(inverse).

Here $e : G \to G$ is the constant map $e(g) = e$ for all $g \in G$. $(e, 1)$ means the map such that $(e, 1)(g) = (e, g)$, etc. A group G is called *commutative* or *Abelian group* if in addition the following diagram commutes

$$\begin{tikzcd}
G \times G \arrow[rr, "T"] \arrow[dr, "\mu"'] & & G \times G \arrow[dl, "\mu"] \\
 & G &
\end{tikzcd}$$

where $T : G \times G \to G \times G$ is the switch map $T(g_1, g_2) = (g_2, g_1)$, for all $(g_1, g_2) \in G \times G$.

A group G *acts* (on the left) on a set A if there is a function $\alpha : G \times A \to A$ such that the following diagrams commute [59]:

1.

$$\begin{tikzcd}
A \arrow[r, "(e, 1)"] \arrow[dr, "1"'] & G \times A \arrow[d, "\alpha"] \\
 & A
\end{tikzcd}$$

2.

$$\begin{tikzcd}
G \times G \times A \arrow[r, "1 \times \alpha"] \arrow[d, "\mu \times 1"'] & G \times A \arrow[d, "\alpha"] \\
G \times A \arrow[r, "\alpha"'] & A
\end{tikzcd}$$

where $(e, 1)(x) = (e, x)$ for all $x \in A$. The *orbits* of the action are the sets $Gx = \{gx : g \in G\}$ for all $x \in A$.

Given two groups $(G, *)$ and $(H, \cdot)$, a *group homomorphism* from $(G, *)$ to $(H, \cdot)$ is a function $h : G \to H$ such that for all x and y in G it holds that

$$h(x * y) = h(x) \cdot h(y).$$

From this property, one can deduce that h maps the identity element e_G of G to the identity element e_H of H, and it also maps inverses to inverses in the sense that $h(x^{-1}) = h(x)^{-1}$. Hence one can say that h is *compatible* with the *group structure.*

The *kernel* $\operatorname{Ker} h$ of a group homomorphism $h : G \to H$ consists of all those elements of G which are sent by h to the identity element e_H of H, i.e.,

$$\operatorname{Ker} h = \{x \in G : h(x) = e_H\}.$$

The *image* $\operatorname{Im} h$ of a group homomorphism $h : G \to H$ consists of all elements of G which are sent by h to H, i.e.,

$$\operatorname{Im} h = \{h(x) : x \in G\}.$$

The kernel is a *normal subgroup* of G and the image is a *subgroup* of H. The homomorphism h is *injective* (and called a *group monomorphism*) iff $\operatorname{Ker} h = e_G$, i.e., iff the kernel of h consists of the identity element of G only.

Similarly, a *ring* (the term introduced by *David Hilbert*) is a set S together with two binary operators $+$ and $*$ (commonly interpreted as addition and multiplication, respectively) satisfying the following conditions:

1. Additive associativity: For all $a, b, c \in S$, $(a + b) + c = a + (b + c)$,
2. Additive commutativity: For all $a, b \in S$, $a + b = b + a$,
3. Additive identity: There exists an element $0 \in S$ such that for all $a \in S$, $0 + a = a + 0 = a$,
4. Additive inverse: For every $a \in S$ there exists $-a \in S$ such that $a+(-a) = (-a) + a = 0$,
5. Multiplicative associativity: For all $a, b, c \in S$, $(a * b) * c = a * (b * c)$,
6. Left and right distributivity: For all $a, b, c \in S$, $a*(b+c) = (a*b)+(a*c)$ and $(b + c) * a = (b * a) + (c * a)$.

A ring is therefore an Abelian group under addition and a semigroup under multiplication. A ring that is commutative under multiplication, has a unit element, and has no divisors of zero is called an *integral domain.* A ring which is also a commutative multiplication group is called a *field.* The simplest rings are the integers $\mathbb{Z}$, polynomials $R[x]$ and $R[x, y]$ in one and two variables, and square $n \times n$ real matrices.

An *ideal* is a subset $\mathfrak{I}$ of elements in a ring R which forms an additive group and has the property that, whenever x belongs to R and y belongs to $\mathfrak{I}$, then xy and yx belong to $\mathfrak{I}$. For example, the set of even integers is an ideal in the ring of integers $\mathbb{Z}$. Given an ideal $\mathfrak{I}$, it is possible to define a factor ring $R/\mathfrak{I}$.

A ring is called *left* (respectively, *right*) *Noetherian* if it does not contain an infinite ascending chain of left (respectively, right) ideals. In this case, the ring in question is said to satisfy the ascending chain condition on left (respectively, right) ideals. A *ring* is said to be *Noetherian* if it is both left and right Noetherian. If a ring R is Noetherian, then the following are equivalent:

1. R satisfies the ascending chain condition on ideals.
2. Every ideal of R is finitely generated.
3. Every set of ideals contains a maximal element.

A *module* is a mathematical object in which things can be added together commutatively by multiplying coefficients and in which most of the rules of manipulating vectors hold. A module is abstractly very similar to a vector space, although in modules, coefficients are taken in rings which are much more general algebraic objects than the fields used in vector spaces. A module taking its coefficients in a ring R is called a module over R or $R-$module. Modules are the basic tool of homological algebra.

Examples of modules include the set of integers $\mathbb{Z}$, the cubic lattice in d dimensions $\mathbb{Z}^d$, and the group ring of a group. $\mathbb{Z}$ is a module over itself. It is closed under addition and subtraction. Numbers of the form $n\alpha$ for $n \in \mathbb{Z}$ and α a fixed integer form a submodule since, for $(n, m) \in \mathbb{Z}$, $n\alpha \pm m\alpha = (n \pm m)\alpha$ and $(n \pm m)$ is still in $\mathbb{Z}$. Also, given two integers a and b, the smallest module containing a and b is the module for their greatest common divisor, $\alpha = GCD(a, b)$.

A module M is a *Noetherian module* if it obeys the ascending chain condition with respect to inclusion, i.e., if every set of increasing sequences of submodules eventually becomes constant. If a module M is Noetherian, then the following are equivalent:

1. M satisfies the ascending chain condition on submodules.
2. Every submodule of M is finitely generated.
3. Every set of submodules of M contains a maximal element.

Let I be a partially ordered set. A *direct system* of *$R-$modules* over I is an ordered pair $\{M_i, \varphi^i_j\}$ consisting of an indexed family of modules $\{M_i : i \in I\}$ together with a family of homomorphisms $\{\varphi^i_j : M_i \to M_j\}$ for $i \leq j$, such that $\varphi^i_i = 1_{M_i}$ for all i and such that the following diagram commutes whenever $i \leq j \leq k$

$$\begin{tikzcd}
M_i \arrow[rr, "\varphi^i_k"] \arrow[dr, "\varphi^j_k"'] & & M_k \\
& M_j \arrow[ur, "\varphi^i_j"'] &
\end{tikzcd}$$

Similarly, an *inverse system* of *$R-$modules* over I is an ordered pair $\{M_i, \psi_i^j\}$ consisting of an indexed family of modules $\{M_i : i \in I\}$ together with a family of homomorphisms $\{\psi_i^j : M_j \to M_i\}$ for $i \leq j$, such that $\psi_i^i = 1_{M_i}$ for all i and such that the following diagram commutes whenever $i \leq j \leq k$

$$\begin{tikzcd}
M_k \arrow[rr, "\psi_i^k"] \arrow[dr, "\psi_j^k"'] & & M_i \\
& M_j \arrow[ur, "\psi_i^j"'] &
\end{tikzcd}$$

2.10 Snake Lemma and Tensor Products

Recall that categories and functors originated from algebraic topology. One of the common theorem-proving tools in algebraic topology is the *snake lemma*, which concerns a commutative and exact diagram called a *snake diagram* [53]:

$$\begin{tikzcd}
& M' \arrow[r, "f"] \arrow[d, "d'"] & M \arrow[r, "g"] \arrow[d, "d"] & M'' \arrow[r] \arrow[d, "d''"] & 0 \\
0 \arrow[r] & N' \arrow[r, "f"'] & N \arrow[r, "g"'] & N'' &
\end{tikzcd}$$

Given a snake diagram as above, the map:

$$\delta : \operatorname{Ker} d'' \longrightarrow \operatorname{Coker} d'$$

is well defined and we have the following exact sequence [53]:

$$\operatorname{Ker} d' \longrightarrow \operatorname{Ker} d \longrightarrow \operatorname{Ker} d'' \xrightarrow{\delta} \operatorname{Coker} d' \longrightarrow \operatorname{Coker} d \longrightarrow \operatorname{Coker} d''$$

where the maps besides δ are the natural ones. The *extended snake diagram* includes the following kernels and cokernels:

$$\begin{tikzcd}
 & \operatorname{Ker} d' \arrow[r] \arrow[d] & \operatorname{Ker} d \arrow[r] \arrow[d] & \operatorname{Ker} d'' \arrow[d] & \\
 & M' \arrow[r] \arrow[d] & M \arrow[r] \arrow[d] & M'' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & N' \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & N'' \arrow[d] & \\
 & \operatorname{Coker} d' \arrow[r] & \operatorname{Coker} d \arrow[r] & \operatorname{Coker} d'' &
\end{tikzcd}$$

together with the connection map:

$$\delta : \operatorname{Ker} d'' \longrightarrow \operatorname{Coker} d'.$$

For example, consider a commutative diagram of $R-$modules and homomorphisms such that each row is exact:

$$\begin{tikzcd}
 & M' \arrow[r] \arrow[d, "f"] & M \arrow[r] \arrow[d, "g"] & M'' \arrow[r] \arrow[d, "h"] & 0 \\
0 \arrow[r] & N' \arrow[r] & N \arrow[r] & N'' &
\end{tikzcd}$$

The following assertions about this diagram can be proved [53]:

1. If f, h are monomorphisms, then g is a monomorphism.
2. If f, h are surjective, then g is surjective.
3. Assume in addition that $0 \longrightarrow M' \longrightarrow M$ is exact and that $N \longrightarrow N'' \longrightarrow 0$ is exact. If any two of f, g, h are isomorphisms, then so is the third.

Now, the following conditions are formally equivalent and define the *tensor exact module F*:

1. For every exact sequence

$$E' \longrightarrow E \longrightarrow E''$$

the following sequence is exact:

$$F \otimes E' \longrightarrow F \otimes E \longrightarrow F \otimes E'',$$

where $\otimes$ defines the tensor product operation, which will be used later for modelling crowd behavior in a topos.

2. For every short exact sequence
$$0 \longrightarrow E' \longrightarrow E \longrightarrow E'' \longrightarrow 0$$
the following sequence is exact:
$$0 \longrightarrow F \otimes E' \longrightarrow F \otimes E \longrightarrow F \otimes E'' \longrightarrow 0.$$
3. For every injection $0 \longrightarrow E' \longrightarrow E$ the following sequence is exact:
$$0 \longrightarrow F \otimes E' \longrightarrow F \otimes E.$$

2.11 A Brief on Categorical Logic

Now we are almost ready to embark on our journey into topos theory. Before that, in this subsection we will make a brief excursion into related area of logic in *coherent Cartesian closed categories.*[20]

A *category with multiplication* is a category $\mathcal{C}$ together with a bifunctor $\cdot : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ and a special object I (a propositional constant, see below). In particular, a *category with binary products* is a category with a binary operation (Cartesian product $\times$) on objects, *projection morphisms* [45, 46]

$$\mathbf{k}^1_{A,B} : A \times B \to A, \qquad \mathbf{k}^2_{A,B} : A \times B \to B,$$

and the *pairing operation* on morphisms $\langle\, , \rangle$ given by

$$\frac{f : C \to A \qquad g : C \to B}{\langle f, g\rangle : C \to A \times B}. \tag{4}$$

The morphisms must satisfy the following set of equations:

$$\begin{aligned} \mathbf{k}^1_{A,B} \circ \langle f, g\rangle &= f, & \mathbf{k}^2_{A,B} \circ \langle f, g\rangle &= g, \\ \langle f, g\rangle \circ h &= \langle f \circ h, g \circ h\rangle, & \langle \mathbf{k}^1_{A,B}, \mathbf{k}^2_{A,B}\rangle &= \mathbf{1}_{A\times B}. \end{aligned}$$

A category has a *terminal object* T iff it has the special morphisms: $\mathbf{k}_A : A \to \mathrm{T}$, which satisfy the equation: for $f : A \to \mathrm{T}, \quad f = \mathbf{k}_A$.

A *cartesian category* is a category with binary products and a terminal object.

---

[20] The term "coherence" covers in category theory what from a logical point of view would be called problems of *completeness*, *axiomatizability* and *decidability.* A coherence condition, or coherence theorem expresses the statement that two or more morphisms between two given objects, the existence of which is given or follows from general properties, are equal. As different authors put stress on different things related to coherence, we stick to MacLane's usage of the term in [44], the primordial paper on coherence. Basically, MacLane has shown that monoidal and symmetric monoidal categories are coherent.

In particular, standard equational axiomatization of Cartesian categories (see [47]) is based on the universality of the Cartesian product and uses as primitives the following morphisms: $\mathbf{1}_A : A \to A$, $\boldsymbol{\pi}_{A,B} : A{\times}B \to A$, $\boldsymbol{\pi}'_{A,B} : A{\times}B \to B$ and $\mathbf{k}_A : A \to \mathrm{I}$ for all objects A and B, and a partial binary operation on morphisms (4). The following equations hold:

$$\begin{aligned} f = \mathbf{k}_A, &\quad \text{for every} \quad f : A \to \mathrm{I}; \\ \boldsymbol{\pi}_{A,B}\langle f, g\rangle = f; &\quad \text{for } f : C \to A \text{ and } g : C \to B; \\ \boldsymbol{\pi}'_{A,B}\langle f, g\rangle = g; &\quad \text{for } f : C \to A \text{ and } g : C \to B; \\ \langle \boldsymbol{\pi}_{A,B}h, \boldsymbol{\pi}'_{A,B}h\rangle = h, &\quad \text{for } h : C \to A{\times}B, \end{aligned}$$

together with the standard categorial equations:

$$\text{(cat 1)} \quad \mathbf{1}_B \circ f = f \circ \mathbf{1}_A = f, \qquad\qquad \text{(cat 2)} \quad h \circ (g \circ f) = (h \circ g) \circ f. \quad (5)$$

Now, we can define the propositional language $\mathcal{P}$ as generated from a set of *propositional letters* $\mathcal{L}$ with the nullary connectives, i.e. propositional constants, I and O, and the binary connectives × and +. The fragments $\mathcal{P}_{\times,+,\mathrm{I}}$, $\mathcal{P}_{\times,+}$ etc. of $\mathcal{P}$ are obtained by keeping only those formulae of $\mathcal{P}$ that contain the connectives in the index. For the propositional letters of $\mathcal{P}$, i.e., for the members of $\mathcal{L}$, we use the schematic letters $p, q, \dots, p_1, \dots$, and for the formulae of $\mathcal{P}$, or of its fragments, we use the schematic letters $A, B, \dots, A_1, \dots$ (see [46]).

Next we define inductively the *terms* that will stand for the morphisms of the free bicartesian category $\mathcal{C}$ generated by $\mathcal{L}$. Every term has a *type*, which is a pair (A, B) of formulae of $\mathcal{P}$. That a term f is of type (A, B) is written $f : A \to B$. The *atomic* terms of $\mathcal{C}$ are for every A of $\mathcal{P}$

$$\begin{gathered} \mathbf{1}_A : A \to A, \\ k_A : A \to \mathrm{I}, \qquad\qquad l_A : \mathrm{O} \to A. \end{gathered}$$

The terms $\mathbf{1}_A$ are called *identities*. The other terms of $\mathcal{C}$ are generated with the following operations on terms, which we present by rules so that from the terms in the premises we obtain the terms in the conclusion (using $f, g, \dots, f_1, \dots$ as schematic letters for terms of $\mathcal{C}$) :

$$\begin{gathered} \frac{f : A \to B \qquad g : B \to C}{g \circ f : A \to C}, \\ \frac{f : A \to C}{K^1_B f : A \times B \to C}, \qquad \frac{f : C \to A}{L^1_B f : C \to A + B}, \\ \frac{f : B \to C}{K^2_A f : A \times B \to C}, \qquad \frac{f : C \to B}{L^2_A f : C \to A + B}, \\ \frac{f : C \to A \qquad g : C \to B}{\langle f, g\rangle : C \to A \times B}, \qquad \frac{f : A \to C \qquad g : B \to C}{[f, g] : A + B \to C}. \end{gathered}$$

The category $\mathcal{C}$ has as objects the formulae of $\mathcal{P}$ and as morphisms equivalence classes of terms[21] so that the both (5) and the following equations are satisfied for $i \in \{1,2\}$ [48]

$$\begin{array}{rlrl}
(K1) & g \circ K^i_A f = K^i_A(g \circ f), & (L1) & L^i_A g \circ f = L^i_A(g \circ f), \\
(K2) & K^i_A g \circ \langle f_1, f_2 \rangle = g \circ f_i, & (L2) & [g_1, g_2] \circ L^i_A f = g_i \circ f, \\
(K3) & \langle g_1, g_2 \rangle \circ f = \langle g_1 \circ f, g_2 \circ f \rangle, & (L3) & g \circ [f_1, f_2] = [g \circ f_1, g \circ f_2], \\
(K4) & \langle K^1_B \mathbf{1}_A, K^2_A \mathbf{1}_B \rangle = \mathbf{1}_{A \times B}, & (L4) & [L^1_B \mathbf{1}_A, L^2_A \mathbf{1}_B] = \mathbf{1}_{A+B}, \\
(k) & \text{for } f : A \to \mathrm{I},\ f = k_A, & (l) & \text{for } f : \mathrm{O} \to A,\ f = l_A.
\end{array}$$

For more technical details on categorical logic, an interested reader might consult J. Lambek's categorial proof-theoretical program [49, 50, 51, 47].

3 A Brief on Topos Theory

In this section we present a minimum of necessary details on topos theory and categorical logic for the purpose of modelling a general crowd dynamics. For more details on topos theory, see [54, 29].

3.1 Topoi: Mathematical Universes with Intuitionistic Logic

Every topos τ can be seen as a *mathematical universe*. As a category, a topos τ possesses a number of structures that generalize constructions that are possible in the category **Set**, which comprises sets and functions. That is, in **Set**, we can construct new sets from given ones in several ways: let S, T be two sets, then we can form the cartesian product $S \times T$, the disjoint union $S \amalg T$ and the exponential S^T, the set of all functions from T to S. These constructions turn out to be fundamental and can all be phrased in an abstract, categorical manner, where they are called finite limits, colimits and exponentials, respectively. By definition, a topos τ has all of these. One consequence of the existence of finite limits is that each topos has a *terminal object*, denoted by 1. This is characterized by the property that for any object A in the topos τ, there exists exactly one morphism from A to 1. In **Set**, a one-element set $1 = \{*\}$ is terminal. As **Set** is a topos too, it is precisely the topos which usually plays the role of our mathematical universe, since we construct our mathematical objects starting from sets and functions between them. As a slogan, we have: *A topos τ is a category similar to* **Set** [42].

Now, in order to 'do mathematics', one must also have a logic, including a deductive system. Each topos comes equipped with an *internal logic*, which

---

[21] Equivalence between proofs in intuitionistic logic is axiomatized independently of these diagrams in the *typed lambda calculus* and in various sorts of categories, like bicartesian closed categories. There, proofs are coded by typed lambda terms or by arrow terms, and two proofs are considered equivalent iff the coding terms are equal as lambda terms or as arrow terms in categories [48, 46].

is of *intuitionistic* type. Recall from Introduction that intuitionistic logic is similar to Boolean logic, the main difference being that the *law of excluded middle* need not hold. In intuitionistic logic, there is a *no* axiom

$$\vdash a \vee \neg a \tag{$*$}$$

like in Boolean logic. Here, $\neg a$ is the negation of the formula (or proposition) a. The algebraic structures representing intuitionistic logic are *Heyting algebras.* Heyting algebra are most simply defined as a certain type of *lattice*, that is a partially-ordered set, or *poset*, $(L \leq)$, in which every pair of elements (x, y) has a least upper bound denoted by $x \vee y$, and a greatest lower bound denoted by $x \wedge y$ (see, e.g. [32]). A top (bottom) element of a lattice L is an element, denoted by 1 (0) such that $x \leq 1$ $(0 \leq x)$ for all $x \in L$. Now a Heyting algebra is defined to be a lattice $(L \leq)$, possessing distinct top and bottom elements, such that, for any pair of elements $(x, y) \in L$, the set of $z \in L$ satisfying $z \wedge x \leq y$ has a largest element. This element, which is uniquely determined by x and y, is denoted by $x \Rightarrow y$. Thus $x \Rightarrow y$ is characterized by the following condition: for all $z \in L$, $z \leq x \Rightarrow y$ iff $z \wedge x \leq y$.

The binary operation on a Heyting algebra which sends each pair of elements $(x, y$ to the element $x \Rightarrow y$ is *implication*; the operation which sends each element x to the element $\neg x = x \Rightarrow 0$ is *negation* (note that the negation satisfies: $z \leq -x$ iff $z \wedge x = 0$ iff $x \leq -z$.)

Heyting algebra is a distributive lattice, i.e., the following equalities hold:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \qquad x \vee (y \wedge z) = (x \vee y) \wedge (x \wedge z).$$

If X is a topological space, then the partially ordered set $(\mathcal{O}(X), \subseteq)$ is a Heyting algebra, where $\mathcal{O}(X)$ is the family of all open sets in X, and $\subseteq$ is the partial ordering of set inclusion. In $\mathcal{O}(X)$, a greatest lower bound and a least upper bound are just set-theoretic intersection and union, while the implication operation is given by: $U \Rightarrow V =$ interior of $(X - U) \cup V$. For any Heyting algebra L, there is a topological space X such that L is isomorphic to a sub-Heyting algebra of $\mathcal{O}(X)$.

Recall from Introduction, that Heyting algebras are associated with theories in *intuitionistic logic*[22] in the same way as Boolean algebras are associated with theories in classical logic. Given a *consistent theory* in an intuitionistic propositional or first-order language L, we can define the equivalence relation $\approx$ on the set of formulas of L by setting $\varphi \approx \psi$ if $T \vdash \varphi \leftrightarrow \psi$. For each

---

[22] Superficially, intuitionistic logic closely resembles propositional and first-order logic. However, intuitionistic operators are not definable in terms of one another in the same way as classical logics, and intuitionistic logic is weaker in the sense that many tautologies of classical logic do not hold in their intuitionistic counterparts. Examples of this include the laws of excluded middle, double negation, elimination and Pierce's law: $((p \Rightarrow q) \Rightarrow p) \Rightarrow p$. Compared with their classical counterparts, intuitionistic logics also display certain kinds of asymmetries, for instance double negation can be introduced but not eliminated.

formula φ we write $[\varphi]$ for its $\approx$-*equivalence class*. Now we define the relation '$\leq$' on the set $H(T)$ of $\approx$-equivalence classes by $[\varphi] \leq [\psi]$ iff $T \vdash \varphi \to \psi$. Then $\leq$ is a partial ordering of $H(T)$ and the partially ordered set $(H(T), \leq)$ is a Heyting algebra in which $[\varphi] \Rightarrow [\psi] = [\varphi \to \psi]$, with analogous equalities defining a greatest lower bound and a least upper bound operations, 0 and 1. $H(T)$ is called the the Heyting algebra determined by T. Heyting algebras of the form $H(T)$ are typical in the sense that, for any Heyting algebra L, there is a propositional intuitionistic theory T such that L is isomorphic to $H(T)$ (see, e.g. [33]).

Here we also need to introduce the fundamental concepts of 'interpretation' and 'model'. An *interpretation* in logic is a possible assignment of semantic meaning, i.e., a total mapping from our variables (or relations, or atoms) – to the elements of $\{false, true\}$. Then: a *model* of a proposition or statement is an interpretation that makes the statement evaluate to *true*.

Obviously, Boolean logic is a special case of intuitionistic logic. It is known from Stone's theorem [43] that each Boolean algebra is isomorphic to an algebra of (clopen, i.e., closed and open) subsets of a suitable topological space.

Let X be a set, and let $P(X)$ be the power set of X, that is, the set of subsets of X. Given a subset $S \in P(X)$, one can ask for each point $x \in X$ whether it lies in S or not. This can be expressed by the *characteristic function* $\chi_S : X \to \{0, 1\}$, which is (for all $x \in X$) defined as

$$\chi_S(x) := \begin{cases} 1, \text{ if } & x \in S, \\ 0, \text{ if } & x \notin S. \end{cases}$$

The two-element set $\{0, 1\}$ plays the role of a set of *truth-values* for propositions (of the form '$x \in S$'). Clearly, 1 corresponds to 'true', 0 corresponds to 'false', and there are no other possibilities. This is an argument about sets, so it takes place in and uses the logic of the topos **Set** of sets and functions. **Set** is a *Boolean topos*, in which the familiar two-valued logic and the axiom $(*)$ hold.[23]

In an arbitrary topos, there is a special object Ω, called the *sub-object classifier*, that takes the role of the set $\{0, 1\} \simeq \{false, true\}$ of truth-values. Let B be an object in the topos, and let A be a sub-object of B. This means that there is a monic $A \to B$,[24] generalising the inclusion of a subset S into a larger set X. Like in **Set**, we can also characterise A as a sub-object of B by a morphism from B to the sub-object classifier Ω.[25] Intuitively, this characteristic morphism from B to Ω tells us how A 'lies in' B. More formally, in a category $\mathcal{C}$ with finite limits, a sub-object classifier is an object Ω, together with a monic (or, monomorphism) *true* $: 1 \to \Omega$, such that to every monic

---

[23] This does not contradict the fact that the internal logic of topoi is intuitionistic, since Boolean logic is a special case of intuitionistic logic.

[24] A *monic*, or monomorphism, is the categorical version of an injective function. In the topos **Set**, monics exactly are injective functions.

[25] In **Set**, this morphism is the characteristic function $\chi_S : X \to \{0, 1\}$.

$m : A \to B$ in $\mathcal{C}$ there is a unique morphism χ which, with the given monic, forms a pullback square:

$$\begin{tikzcd}
A \arrow[r] \arrow[d, "m"'] & 1 \arrow[d, "true"] \\
B \arrow[r, "\chi"'] & \Omega
\end{tikzcd}$$

In **Set**, the morphism true : $1 \longrightarrow \{0, 1\}$ is given by true$(*) = 1$. In general, the sub-object classifier Ω need not be a set, since it is an object in the topos τ, and the objects of τ need not be sets. Nonetheless, there is an abstract notion of *elements* (or *points*) in category theory that we can use. The elements of Ω are the truth-values available in the internal logic of our topos τ, just like 'false' and 'true', the elements of $\{false, true\}$, are the truth-values available in the topos **Set**.

To understand the abstract notion of elements, let us consider sets for a moment. Let $1 = \{*\}$ be a one-element set (or, a singleton), the terminal object in **Set**. Let S be a set and consider a morphism $e : 1 \longrightarrow S$, from 1 to S. Clearly, $e(*) \in S$ is one element of S. The set of all functions from 1 to S corresponds exactly to the elements of S. This idea can be generalised to other categories: if there is a terminal object 1, then we consider morphisms from 1 to an object A in the category as *elements of* A. For example, in the definition of the sub-object classifier Ω, the morphism $true : 1 \to \Omega$ is an element of Ω. It may happen that an object A has no elements, i.e., there are no morphisms $1 \to A$.

As mentioned, the elements of the sub-object classifier, understood as the morphisms $1 \to \Omega$, are the truth-values. Moreover, the set of these morphisms forms a Heyting algebra (see, e.g. [54]). This is how (the algebraic representation of) intuitionistic logic manifests itself in a topos. Another, closely related fact is that the sub-objects of any object A in a topos form a Heyting algebra [42].

3.2 The Topos of Varying Sets

In standard set theory, to each subset A of a set X there is associated a *characteristic map* $\chi^A : X \to \{0, 1\}$ defined by:

$$\chi^A(x) := \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

so that

$$A = (\chi^A)^{-1}\{1\}. \tag{7}$$

Conversely, any function $\chi : X \to \{0,1\}$ defines a unique subset $A_\chi := \chi^{-1}\{1\}$ of X whose characteristic function is equal to χ.

Next, consider a *hypothetical crowd mechanics* whose basic ingredient is a Boolean lattice L of propositions about the *crowd space-time universe.* A 'pure state' σ of the system will give rise to a *valuation* on L, i.e., a homomorphism $V^\sigma : L \to \Omega$ from L to the simplest Boolean algebra $\Omega := \{0,1\}$ with '0' interpreted as $false$ and '1' as $true$. Thus a L-valuation, or L-model, is a characteristic map that is also a homomorphism between Boolean algebras [34].

Now let us consider what a probabilistic version of such a theory might look like.[26] Consider the proposition "$\alpha \in L$ is true with probability p" (to be denoted by $\langle \alpha, p\rangle$) is to be read as saying that the state of affairs represented by α has an 'intrinsic tendency' to occur that is measured by the number $p \in [0,1]$. In other words, "the probability of the truthfulness of α is p" – might be an interpretation of this. Note that to each probability measure μ on L (a 'statistical state' of the system), and for each $p \in [0,1]$, there is associated the subset of all $\alpha \in L$ such that $\mu(\alpha) = p$. In turn, this gives rise to the characteristic map $\chi^{\mu,p} : L \to \{0,1\}$, defined by

$$\chi^{\mu,p}(\alpha) := \begin{cases} 1, & \text{if } \mu(\alpha) = p \ ; \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

as a particular example of the situation represented by (6).

We can think of the *second-level propositions* $\langle \alpha, p\rangle$ as generating a new logical algebra with respect to which each measure μ on L produces a genuine $\{0,1\}$-valued valuation, or a $\{0,1\}$-valued model V^μ, defined by [34]

$$V^\mu\langle \alpha, p\rangle := \begin{cases} 1, & \text{if } \mu(\alpha) = p \ ; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Thus, for example, the conjunction operation on these new propositions is *defined* to be such that, for all μ,

$$V^\mu(\langle \alpha, p\rangle \wedge \langle \beta, q\rangle) := \begin{cases} 1, & \text{if } \mu(\alpha) = p \ \text{ and } \ \mu(\beta) = q; \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

This leads naturally to the idea of two second-level propositions being *μ-semantically equivalent* if their V^μ valuations are equal, and *semantically*

---

[26] We are heading here towards an intuitionistic probability interpretation rather then classical Bayesian probability interpretation. This approach could be described as 'intuitionistic doxastic' if we were to extend with modal operators ∇ and $\Box$ in addition to the classical connectives. Note that *doxastic logic* is a a *modal logic* concerned with reasoning about beliefs. Also, note that *intuitionistic logic* can be defined in a variety of ways, including Gentzen-style sequent and sequent-like calculi, Hilbert-style axiomatic calculi, and natural deduction systems.

equivalent if they are μ-semantically equivalent for all measures μ. For example, for all μ and all $p \in [0,1]$ we have

$$V^{\mu}\langle \alpha, p\rangle = V^{\mu}\langle \neg\alpha, 1-p\rangle, \tag{11}$$

since $\mu(\alpha) + \mu(\neg\alpha) = 1$ for all $\alpha \in L$. Hence $\langle \alpha, p\rangle$ and $\langle \neg\alpha, 1-p\rangle$ are semantically equivalent for all $p \in [0,1]$. A more complex example is given by the result that, for any disjoint propositions α and β (i.e., $\alpha \wedge \beta = 0$), we have [34]

$$V^{\mu}\langle \alpha \vee \beta, p\rangle = V^{\mu}\Big(\bigvee_{q\in[0,1]} \langle \alpha, p-q\rangle \wedge \langle \beta, q\rangle\Big), \tag{12}$$

which arises from the fact that $\mu(\alpha \vee \beta) = \mu(\alpha) + \mu(\beta)$ for any such pair of propositions. Thus we see that, if $\alpha\wedge\beta = 0$, then $\langle \alpha\vee\beta, p\rangle$ and $\bigvee_{q\in[0,1]}\langle \alpha, p-q\rangle \wedge \langle \beta, q\rangle$ are semantically equivalent for all $p \in [0,1]$.

3.3 Sets through Time

As an example of how contexts and generalised semantic values can arise, consider a fixed set X of people who are all alive at some initial time, and whose bodies are preserved once they die (and who are still referred to as 'people' in that state). Thus if $D(t) \subseteq X$ denotes the subset of dead people at any time t, then as t increases $D(t)$ will clearly stay constant or increase, i.e., $t_1 \leq t_2$ implies $D(t_1) \subseteq D(t_2)$. Such a parameterised family of sets $D(t)$, $t \in \mathbb{R}$, is an example of what has been called a "set through time"[27] by those working in the foundations of topos theory [27, 28, 29].

Now suppose that some members of our population are actually immortal. Then what truth value should be assigned to the proposition 'person x is mortal' if all truth statements are required to be verifiable in some operational sense? If x is alive, the proposition cannot be said to be true, on the assumption that mortality of a living being cannot be verified operationally, but neither can it be denied, since even if x *is* numbered among the immortals there is no way of showing this. Thus we are lead to the notion of a 'stage of truth' as the context in which a proposition acquires meaning, in our case, the time t, and to the idea that the truth values of a statement at a stage t may not just lie in the set $\{0,1\}$.

Topos theory provides an answer to this problem that stems from the observation that there may be a later time t at which x *does* die, and then of course $x \in D(t')$ for all times $t' \geq t$. A key idea in the theory of sets-through-time is that the 'truth value', or 'semantic value' at the stage t_0 of the proposition 'x is mortal' is defined to be the set $\chi^D_{t_0}(x)$ of all later times t at which x is dead:

$$\chi^D_{t_0}(x) := \{\, t \geq t_0 \mid x \in D(t)\,\}. \tag{13}$$

---

[27] A 'set through time' can be interpreted as 'temporal', which uses a model of time (linear, branching, etc.) or as 'dynamic', using change without time.

Note that if x never dies, i.e., if he or she is immortal, then the right hand side of (13) is just the empty set. On the other hand, x is dead at a time t iff

$$\chi_t^D(x) = \uparrow(t) := [t, \infty). \tag{14}$$

Equivalently, at stage t we have

$$D(t) = (\chi_t^D)^{-1}\{\uparrow(t)\}. \tag{15}$$

When compared with (6), the relation (15) shows that the parameterised family of maps $\chi_{t_0}^D : X \to \Omega(t_0)$, $t_0 \in \mathbb{R}$, (where $\Omega(t_0)$ denotes the collection of all upper sets lying above t_0) is the analogue of the single characteristic function of normal set theory.

From a logical perspective, the crucial property of this set $\Omega(t_0)$ of all possible semantic values at stage t_0 is that it possesses the structure of a Heyting algebra. Thus $\Omega(t_0)$ is a distributive lattice with the property that for any $a, b \in \Omega(t_0)$ there is a unique element $(a \Rightarrow b) \in \Omega(t_0)$ (with the intuitive meaning "if a *then* b") satisfying

$$c \leq (a \Rightarrow b) \quad \text{iff} \quad c \wedge a \leq b. \tag{16}$$

The negation operation in such an algebra is defined by $\neg a := (a \Rightarrow 0)$, and satisfies the relation $a \leq \neg\neg a$ for all a.[28] Indeed, it can be shown that a Heyting algebra is Boolean iff $a = \neg\neg a$ for all a [34].

3.4 Presheaves on a Poset

The ideas sketched above extend readily to the situation where the 'stages of truth' are elements of a general partially-ordered set (or *poset* for short, see, e.g. [28]). The necessary mathematical development is most naturally expressed in the language of category theory.

In the special case of the category of sets, **Set**, the objects are sets and a morphism is a function between a pair of sets. In general, each morphism f in a category is associated with a pair of objects, known as its 'domain' and 'codomain', and is written as $f : B \to A$ where B and A are the domain and codomain respectively. In the case of the category **Set**, this is just the usual composition of functions.

A simple example of a category is given by any poset $\mathcal{P}$:

1. The objects are defined to be the elements of $\mathcal{P}$; and
2. If $p, q \in \mathcal{P}$, a morphism from p to q is defined to exist iff $p \preceq q$ in the poset structure.

---

[28] This is one of the main reasons why a Heyting algebra is chosen as the formal mathematical structure that underlies intuitionistic logic. Thus there is a strong connection between the theory of sets through time and the logic of intuitionism.

Thus, in a poset regarded as a category, there is at most one morphism between any pair of objects $p, q \in \mathcal{P}$; if it exists, we will write this morphism as $i_{pq} : p \rightarrow q$. This example will be important for us in form of the *category of contexts.*

From our perspective, the most relevant feature of a topos τ is that it is a category which behaves in many ways like the category of sets **Set** [54, 29]. Here, we will list its most important properties (for our purposes):

1. There is a *terminal object* 1_τ in τ; this means that given any object A in the topos, there is a unique morphism $A \rightarrow 1_\tau$.
   For any object A in the topos τ, a morphism $1_\tau \rightarrow A$ is called a *global element*[29] of A. The set of all global elements of A is denoted $\mathit{\Gamma} A$.
   Given $A, B \in \mathrm{Ob}(\tau)$, there is a product $A \times B$ in τ. In fact, a topos always has *pull-backs*, and the product is just a special case of this.[30]
2. There is an *initial object* 0_τ in τ. This means that given any object A in the topos, there is a unique morphism $0_\tau \rightarrow A$.
   Given $A, B \in \mathrm{Ob}(\tau)$, there is a co-product $A \sqcup B$ in τ. In fact, a topos always has *push-outs*, and the co-product is just a special case of this.[31]
3. There is *exponentiation*: i.e., given objects A, B in τ we can form the object A^B, which is the topos analogue of the set of functions from B to A in set theory. The definitive property of exponentiation is that, given any object C, there is an isomorphism

$$\mathtt{Mor}_\tau C A^B \simeq \mathtt{Mor}_\tau C \times B A \tag{17}$$

   that is natural in A and C.
4. There is a sub-object classifier $\mathit{\Omega}_\tau$.

The last item is of particular importance to us as it is the source of the Heyting algebras. To explain what is meant, let us first consider the familiar topos, **Set**, of sets. There, the subsets $K \subseteq X$ of a set X are in one-to-one correspondence with functions $\chi_K : X \rightarrow \{0, 1\}$, where $\chi_K(x) = 1$ if $x \in K$, and $\chi_K(x) = 0$ otherwise. Thus the target space $\{0, 1\}$ can be regarded as the simplest 'false-true' Boolean algebra, and the mathematical proposition "$x \in K$" is true if $\chi_K(x) = 1$, and false otherwise.

In the case of a topos τ, the sub-objects[32] K of an object X in the topos are in one-to-one correspondence with morphisms $\chi_K : X \rightarrow \mathit{\Omega}_\tau$, where the special object $\mathit{\Omega}_\tau$ is the *sub-object classifier* (or 'object of truth values'), and

---

[29] In the category of sets, **Set**, the terminal object $1_{\mathbf{Set}}$ is a singleton set $\{*\}$. It follows that the elements of $\mathit{\Gamma} A$ are in one-to-one correspondence with the elements of A.

[30] The conditions in 1. above are equivalent to saying that τ is *finitely complete.*

[31] The conditions in 2. above are equivalent to saying that τ is *finitely co-complete.*

[32] An object K is a *sub-object* of another object X if there is a monic morphism $K \hookrightarrow X$. In the topos **Set** of sets, this is equivalent to saying that K is a subset of X.

plays an analogous role to that of $\{0,1\}$ in the category of sets. An important property for us is that, in any topos τ, the collection, $\mathrm{Sub}(A)$, of sub-objects of an object A forms a *Heyting algebra*. The reader is referred to the standard texts for proofs (e.g., see [54]).

The idea of a presheaf. To illustrate the main ideas, we will first give a few definitions from the theory of presheaves on a partially-ordered set (or, *poset*). A *presheaf* (also known as a *varying set*) X on a poset $\mathcal{P}$ is a function that assigns to each $p \in \mathcal{P}$, a set X_p, and to each pair $p \preceq q$ (i.e., $i_{pq} : p \to q$), a map $X_{qp} : X_q \to X_p$ such that (i) $X_{pp} : X_p \to X_p$ is the identity map id_{X_p} on X_p, and (ii) whenever $p \preceq q \preceq r$, the composite map $X_r \xrightarrow{X_{rq}} X_q \xrightarrow{X_{qp}} X_p$ is equal to $X_r \xrightarrow{X_{rp}} X_p$, so that

$$X_{rp} = X_{qp} \circ X_{rq}. \tag{18}$$

The notation X_{qp} is shorthand for the more cumbersome $X(i_{pq})$; see below in the definition of a functor.

A *natural transformation* $\eta : X \to Y$ between two presheaves X, Y on $\mathcal{P}$ is a family of maps $\eta_p : X_p \to Y_p$, $p \in \mathcal{P}$, that satisfy the intertwining conditions

$$\eta_p \circ X_{qp} = Y_{qp} \circ \eta_q\,, \tag{19}$$

whenever $p \preceq q$. In other words, we have the commutative diagram:

$$\begin{tikzcd} X_q \arrow[r, "X_{qp}"] \arrow[d, "\eta_q"'] & X_p \arrow[d, "\eta_p"] \\ Y_q \arrow[r, "Y_{qp}"'] & Y_p \end{tikzcd} \tag{20}$$

A *sub-object* of a presheaf X is a presheaf K, with a morphism $i : K \to X$ such that (i) $K_p \subseteq X_p$ for all $p \in \mathcal{P}$; and (ii) for all $p \preceq q$, the map $K_{qp} : K_q \to K_p$ is the restriction of $X_{qp} : X_q \to X_p$ to the subset $K_q \subseteq X_q$. This is shown in the commutative diagram

$$\begin{tikzcd} K_q \arrow[r, "K_{qp}"] \arrow[d] & K_p \arrow[d] \\ X_q \arrow[r, "X_{qp}"'] & X_p \end{tikzcd} \tag{21}$$

where the vertical morphisms are subset inclusions (see [54, 29]).

A simple, but important, special case is when the varying set $X(p)$ is constant i.e., $X(p) = X$ for all $p \in \mathcal{P}$, and X_{pq} is the identity map from $X = X(p)$ to $X = X(q)$ for all pairs $p \leq q$. In this situation, each set $A(p)$, $p \in \mathcal{P}$, can be regarded as a subset of the fixed set X, and the condition on a varying set $A := \{A(p), p \in \mathcal{P}\}$ to be a sub-object of X is simply that

$$p \leq q \text{ implies } A(p) \subseteq A(q). \tag{22}$$

This special case where $X(p)$ is constant also gives rise to the varying-set analogue of a 'complement' of a subset. The obvious family of subsets of X to serve as the complement of $\{A(p), p \in \mathcal{P}\}$ is $\{X - A(p), p \in \mathcal{P}\}$, but this does not give a proper varying set since $p \leq q$ implies $X - A(p) \supseteq X - A(q)$, which is the wrong behavior. The appropriate definition is the genuine varying set $\neg A := \{\neg A(p), p \in \mathcal{P}\}$, where

$$\neg A(p) := \{x \in X \mid \forall q \geq p, x \notin A(q)\}. \tag{23}$$

It follows that $x \notin \neg A(p)$ iff there is some $q \geq p$ such that $x \in A(q)$, and hence

$$\neg\neg A(p) := \{x \in X \mid \forall q \geq p \, \exists \, r \geq q \text{ s.t. } x \in A(r)\}. \tag{24}$$

We see that $A(p) \subseteq \neg\neg A(p)$ whereas, in normal set theory, the double complement of a subset is always equal to the subset itself. This non-standard behavior in the varying-set theory is a reflection of the fact that the underlying logical structure is non-Boolean.

As in the case of sets through time, a key role is played by the collections $\Omega(p)$, $p \in \mathcal{P}$, of all upper sets lying above p. More precisely, a *sieve*[33] on p in $\mathcal{P}$ is defined to be any subset S of $\mathcal{P}$ such that if $r \in S$ then (i) $r \geq p$, and (ii) $r' \in S$ for all $r' \geq r$. For each $p \in \mathcal{P}$, the set $\Omega(p)$ of all sieves on p can be shown to be a *Heyting algebra*, and for all pairs $p \leq q$ there is a natural map $\Omega_{pq} : \Omega(p) \to \Omega(q)$ defined by [34]

$$\Omega_{pq}(S) := S \cap \uparrow(q), \tag{25}$$

where $\uparrow(q) := \{r \in \mathcal{P} \mid r \geq q\}$ is the unit element in the Heyting algebra $\Omega(q)$ (the null element is the empty set). It is easy to see that, with the maps Ω_{pq} in (25), $\Omega := \{\Omega(p), p \in \mathcal{P}\}$ is a varying set over $\mathcal{P}$ and hence an object in the category $\mathbf{Set}^{\mathcal{P}}$.

A very important example of the use of Ω occurs if A is a sub-object of the object X. There is then an associated *characteristic* morphism $\chi^A : X \to \Omega$ with, at each stage $p \in \mathcal{P}$, the 'component' $\chi^A_p : X(p) \to \Omega(p)$ being defined by

$$\chi^A_p(x) := \{\, q \geq p \mid X_{pq}(x) \in A(q) \,\}, \tag{26}$$

where the fact that the right hand side of (26) actually *is* a sieve on p in $\mathcal{P}$ follows from the defining properties of a sub-object. Thus in each 'branch'

---

[33] This is the notation employed by in [28]; other authors (for example, [29] use the term 'cosieve' for what Bell calls a 'sieve', and *vice versa*.

of the poset going up from p, $\chi_p^A(x)$ picks out the first member q (the "time till truth") in that branch for which $X_{pq}(x)$ lies in the subset $A(q)$, and the commutative diagram (20) then guarantees that $X_{pr}(x)$ will lie in $A(r)$ for all $r \geq q$. In the special case where $X(p) = X$ for all p, (27) simplifies to [34]

$$\chi_p^A(x) := \{\, q \geq p \mid x \in A(q) \,\}. \tag{27}$$

In what follows, the expression (27) plays a crucial role as the analogue in the theory of varying sets of the characteristic map (6) $\chi^A : X \to \{0, 1\}$ of normal set theory. Indeed, the analogue of the relation (7) for this situation is at each stage $p \in \mathcal{P}$ given by

$$A(p) = (\chi_p^A)^{-1}\{\, \uparrow\!(p) \,\}. \tag{28}$$

Conversely, each morphism $\chi : X \to \Omega$ defines a sub-object of X (via (28)), and for this reason the object Ω in $\mathbf{Set}^{\mathcal{P}}$ is known as the *sub-object classifier* in the category $\mathbf{Set}^{\mathcal{P}}$; the existence of such an object is one of the defining properties[34] for a category to be a topos, so $\mathbf{Set}^{\mathcal{P}}$ must have such an object. As the target of characteristic maps (i.e., the analogue of $\{0, 1\}$ in normal set theory), Ω can be thought of as the 'object of truth values'—an assignation that is reinforced by the observation that Ω has an internal structure of a Heyting algebra. For example, the conjunction $\wedge : \Omega \times \Omega \to \Omega$ is defined to be the morphism in the category $\mathbf{Set}^{\mathcal{P}}$ whose components $\wedge_p : \Omega(p) \times \Omega(p) \to \Omega(p)$, $p \in \mathcal{P}$, are the conjunctions $\wedge_p$ in the 'local' Heyting algebras $\Omega(p)$; the other logical operations are defined in a similar way.

3.5 Presheaves on a General Category

The ideas sketched above admit an immediate generalization to the theory of presheaves on an arbitrary 'small' category $\mathcal{C}$[35] Here, a central idea is that *a presheaf is a functor*, a structure-preserving map between a pair of categories $\mathcal{C}$ and $\mathcal{D}$.

As a poset $\mathcal{P}$ is itself a category, then a presheaf on the poset $\mathcal{P}$ is a contravariant functor $\mathcal{X} : \mathcal{P} \longrightarrow \mathbf{Set}$, from the category $\mathcal{P}$ to the category $\mathbf{Set}$ of normal sets. Equivalently, it is a covariant functor $\mathcal{X} : \mathcal{P}^{\mathrm{op}} \longrightarrow \mathbf{Set}$, from the 'opposite' category $\mathcal{P}^{\mathrm{op}}$ to $\mathbf{Set}$.[36]

Presheaves on an arbitrary category $\mathcal{C}$. These remarks motivate the definition of a presheaf on an arbitrary small category $\mathcal{C}$: namely, a *presheaf*

---

[34] Another defining property for a category $\mathcal{P}$ to be a topos is that a Cartesian product $A \times B$ exists for any pair of objects A, B in $\mathcal{P}$. For the full definition see one of the standard texts (e.g., [28, 29].

[35] The qualification 'small' means that the collection of objects is a genuine set, as is the collection of all morphisms between any pair of objects.

[36] Note that categoricians usually call the objects in $\mathcal{P}$ *stages of truth*, or just 'stages'. Following Chris Isham, we will call them 'contexts'.

on $\mathcal{C}$ is a covariant functor $X : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, or equivalently, a presheaf is a contravariant functor $X : \mathcal{C} \to \mathbf{Set}$.[37]

We want to make the collection of presheaves on $\mathcal{C}$ into a category, and therefore we need to define what is meant by a 'morphism' between two presheaves X and Y. The intuitive idea is that such a morphism from X to Y must give a 'picture' of X within Y. Formally, such a morphism is defined to be a *natural transformation* $N : X \to Y$, by which is meant a family of maps (called the *components* of N) $N_A : X_A \to Y_A$, $A \in \mathrm{Ob}(\mathcal{C})$, such that if $f : B \to A$ is a morphism in $\mathcal{C}$, then the composite map $X_A \xrightarrow{N_A} Y_A \xrightarrow{Y(f)} Y_B$ is equal to $X_A \xrightarrow{X(f)} X_B \xrightarrow{N_B} Y_A$. In other words, we have the commutative diagram [38]

$$\begin{tikzcd}
X_A \arrow[r, "X(f)"] \arrow[d, "N_A"'] & X_B \arrow[d, "N_B"] \\
Y_A \arrow[r, "Y(f)"'] & Y_B
\end{tikzcd} \tag{29}$$

of which (20) is clearly a special case. The category of presheaves on $\mathcal{C}$ equipped with these morphisms is denoted $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$.

The idea of a *sub-object* generalizes in an obvious way. Thus we say that K is a *sub-object* of X if there is a morphism in the category of presheaves (i.e., a natural transformation) $\iota : K \to X$ with the property that, for each A, the component map $\iota_A : K_A \to X_A$ is a subset embedding, i.e., $K_A \subseteq X_A$. Thus, if $f : B \to A$ is any morphism in $\mathcal{C}$, we get the analogue of the commutative diagram (21):

$$\begin{tikzcd}
K_A \arrow[r, "K(f)"] \arrow[d] & K_B \arrow[d] \\
X_A \arrow[r, "X(f)"'] & X_B
\end{tikzcd} \tag{30}$$

where, once again, the vertical morphisms are subset inclusions.

The category $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$ of presheaves on $\mathcal{C}$ forms a topos.

Sieves and the sub-object classifier Ω. Among the key concepts in presheaf theory is that of a 'sieve', which plays a central role in the construction of the sub-object classifier in the topos of presheaves on a category $\mathcal{C}$.

---

[37] For simplicity, from now on, instead of calligraphic letters we will use ordinary Latin letters for functors/presheafs: e.g., X instead of $\mathcal{X}$, etc.

A *sieve* on an object A in $\mathcal{C}$ is defined to be a collection S of morphisms $f : B \to A$ in $\mathcal{C}$ with the property that if $f : B \to A$ belongs to S, and if $g : C \to B$ is any morphism with co-domain B, then $f \circ g : C \to A$ also belongs to S. In the simple case where $\mathcal{C}$ is a poset, a sieve on $p \in \mathcal{C}$ is any subset S of $\mathcal{C}$ such that if $r \in S$ then (i) $r \preceq p$, and (ii) $r' \in S$ for all $r' \preceq r$; in other words, a sieve is nothing but a *lower* set in the poset.

The presheaf $\Omega : \mathcal{C} \to \mathbf{Set}$ is now defined as follows. If A is an object in $\mathcal{C}$, then Ω_A is defined to be the set of all sieves on A; and if $f : B \to A$, then $\Omega(f) : \Omega_A \to \Omega_B$ is defined as

$$\Omega(f)(S) := \{h : C \to B \mid f \circ h \in S\}, \tag{31}$$

for all $S \in \Omega_A$; the sieve $\Omega(f)(S)$ is often written as $f^*(S)$, and is known as the *pull-back* to B of the sieve S on A by the morphism $f : B \to A$.

It should be noted that if S is a sieve on A, and if $f : B \to A$ belongs to S, then from the defining property of a sieve we have

$$f^*(S) := \{h : C \to B \mid f \circ h \in S\} = \{h : C \to B\} =: {\downarrow}B\,, \tag{32}$$

where ${\downarrow}B$ denotes the *principal sieve* on B, defined to be the set of all morphisms in $\mathcal{C}$ whose codomain is B.[38]

If $\mathcal{C}$ is a poset, the pull-back operation corresponds to a family of maps $\Omega_{qp} : \Omega_q \to \Omega_p$ (where Ω_p denotes the set of all sieves/lower sets on p in the poset) defined by $\Omega_{qp} = \Omega(i_{pq})$ if $i_{pq} : p \to q$ (i.e., $p \preceq q$). It is straightforward to check that if $S \in \Omega_q$, then

$$\Omega_{qp}(S) := {\downarrow}p \cap S\,, \tag{33}$$

where ${\downarrow}p := \{r \in \mathcal{C} \mid r \preceq p\}$.

A crucial property of sieves is that the set Ω_A of sieves on A has the structure of a *Heyting algebra*. Specifically, Ω_A is a Heyting algebra where the unit element 1_{Ω_A} in Ω_A is the principal sieve ${\downarrow}A$, and the null element 0_{Ω_A} is the empty sieve $\emptyset$. The partial ordering in Ω_A is defined by $S_1 \preceq S_2$ iff $S_1 \subseteq S_2$; and the logical connectives are defined as [38]:

$$S_1 \land S_2 := S_1 \cap S_2\,, \tag{34}$$

$$S_1 \lor S_2 := S_1 \cup S_2\,, \tag{35}$$

$$S_1 \Rightarrow S_2 := \{f : B \to A \mid \forall\ g : C \to B \text{ if } f \circ g \in S_1 \text{ then } f \circ g \in S_2\}. \tag{36}$$

As in any Heyting algebra, the negation of an element S (called the *pseudo-complement* of S) is defined as $\neg S := S \Rightarrow 0$; so that

$$\neg S := \{f : B \to A \mid \text{for all } g : C \to B,\ f \circ g \notin S\}. \tag{37}$$

It can be shown that the presheaf Ω is a sub-object classifier for the topos $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$. In other words, sub-objects of any object X in this topos (i.e., any

---

[38] In words: the pull-back of any sieve on A by a morphism from B to A that belongs to the sieve, is the principal sieve on B.

presheaf on $\mathcal{C}$) are in one-to-one correspondence with morphisms $\chi : X \to \Omega$. This works as follows. First, let K be a sub-object of X. Then there is an associated *characteristic* morphism $\chi_K : X \to \Omega$, whose 'component' $\chi_{KA} : X_A \to \Omega_A$ at each stage/context A in $\mathcal{C}$ is defined (for all $x \in X_A$) as:

$$\chi_{KA}(x) := \{f : B \to A \mid X(f)(x) \in K_B\}. \tag{38}$$

From the defining properties of a sub-object, it follows that the right hand side of (38) actually *is* a sieve on A.

Thus, in each 'branch' of the category $\mathcal{C}$ going 'down' from the stage A, $\chi_{KA}(x)$ picks out the first member B in that branch for which $X(f)(x)$ lies in the subset K_B, and the commutative diagram (30) then guarantees that $X(h \circ f)(x)$ will lie in K_C for all $h : C \to B$. Thus each stage A in $\mathcal{C}$ serves as a possible context for an assignment to each $x \in X_A$ of a generalised truth value—a sieve belonging to the Heyting algebra Ω_A. This is the sense in which contextual, generalised truth values arise naturally in a topos of presheaves.

There is a converse to (38): namely, each morphism $\chi : X \to \Omega$ (i.e., a natural transformation between the presheaves X and Ω) defines a sub-object K^χ of X at each stage A via

$$K^\chi_A := \chi_A^{-1}\{1_{\Omega_A}\}. \tag{39}$$

Global elements of a presheaf. We recall that, in any topos τ a *terminal object* is defined to be an object 1_τ with the property that, for any object X in the category, there is a unique morphism $X \to 1_\tau$; it is easy to show that terminal objects are unique up to isomorphism. A *global element* of an object X is then defined to be any morphism $s : 1_\tau \to X$. The motivation for this nomenclature is that, in the case of the category of sets, a terminal object is any singleton set $\{*\}$; and then it is true that there is a one-to-one correspondence between the elements of a set X and functions from $\{*\}$ to X.

For the category of presheaves on $\mathcal{C}$, a terminal object $1 : \mathcal{C} \to \mathbf{Set}$ can be defined by $1_A := \{*\}$ at all stages A in $\mathcal{C}$; if $f : B \to A$ is a morphism in $\mathcal{C}$ then $1(f) : \{*\} \to \{*\}$ is defined to be the map $* \mapsto *$. This is indeed a terminal object since, for any presheaf X, we can define a unique natural transformation $N : X \to 1$ whose components $N_A : X(A) \to 1_A = \{*\}$ are the constant maps $x \mapsto *$ for all $x \in X_A$.

A global element of a presheaf X is also called a *global section.* As a morphism $\gamma : 1 \to X$ in the topos $\mathbf{Set}^{\mathcal{C}^{\mathrm{op}}}$, a global element corresponds to a choice of an element $\gamma_A \in X_A$ for each stage A in $\mathcal{C}$, such that, if $f : B \to A$, the following 'matching condition' is satisfied [38]:

$$X(f)(\gamma_A) = \gamma_B. \tag{40}$$

3.6 Classical Realism vs. Quantum Instrumentalism

In *classical physics*, one has a space of states $\mathcal{S}$, and physical quantities A are represented by measurable real-valued functions $f_A : \mathcal{S} \to \mathbb{R}$. A proposition about a physical quantity A is of the form '$A \in \Delta$', which means "the physical quantity A has a value in the *Borel set*[39] Δ." This proposition is represented by the inverse image $f_A^{-1}(\Delta) \subseteq \mathcal{S}$. In general, propositions about the physical system correspond to Borel subsets of the state space $\mathcal{S}$ [38, 39, 40, 41]. If we have two propositions '$A \in \Delta_1$', '$B \in \Delta_2$' and the corresponding subsets $f_A^{-1}(\Delta_1)$, $f_B^{-1}(\Delta_2)$, then the intersection $f_A^{-1}(\Delta_1) \cap f_B^{-1}(\Delta_2)$ corresponds to the proposition '$A \in \Delta_1$ and $B \in \Delta_2$', the union $f_A^{-1}(\Delta_1) \cup f_B^{-1}(\Delta_2)$ corresponds to '$A \in \Delta_1$ or $B \in \Delta_2$', and the complement $\mathcal{S} \backslash f_A^{-1}(\Delta_1)$ corresponds to the negation '$A \notin \Delta_1$'. Moreover, given a state s, i.e., an element of the state space $\mathcal{S}$, each proposition is either true or false: if s lies in the subset of $\mathcal{S}$ representing the proposition, then the proposition is true, otherwise it is false. Every physical quantity A has a value in the state s, namely $f_A(s) \in \mathbb{R}$. Thus classical physics is a *realist* theory in which propositions have *truth-values independent of measurements and observers.* The logic is Boolean, since classical physics is based on constructions with sets and functions, i.e., it takes place in the topos **Set**. We take this as a rule: if we want to describe a physical system S as a classical system, then the topos **Set** is used. This will be our *framework for crowd mechanics.*

On the other hand, in *quantum theory*, the mathematical description is very different [38, 39, 40, 41]. Physical quantities A are represented by self-adjoint operators $\widehat{A}$ on a Hilbert space $\mathcal{H}$. While $\mathcal{H}$ can be called a space of states, the states $\psi \in \mathcal{H}$ play a very different role from those in classical theory. In particular, a state ψ does not assign values to all physical quantities, only to those for which ψ happens to be an *eigenstate*. The spectral theorem shows that propositions '$A \in \Delta$' are represented by projection operators $\widehat{E}[A \in \Delta]$ on Hilbert space. Unless ψ is an eigenstate of A, such a proposition is neither true nor false (except for the trivial cases $\widehat{E}[A \in \Delta] = \widehat{0}$, which represents trivially false propositions, and $\widehat{E}[A \in \Delta] = \widehat{1}$, which represents trivially true propositions). The mathematical formalism of quantum theory is interpreted in an *instrumentalist* manner: given a state ψ, the proposition '$A \in \Delta$' is assigned a probability of being true, given by the expectation value

---

[39] A Borel set is any set in a topological space that can be formed from open sets (or, equivalently, from closed sets) through the operations of countable union, countable intersection, and relative complement. For a topological space X, the collection of all Borel sets on X forms a σ-algebra, known as the Borel σ-algebra; it is the smallest σ-algebra containing all open sets (or, equivalently, all closed sets).

Borel sets are important in Measure theory, which (among other things) is important in probability; i.e., every probability space has a measure that takes values in $[0, 1]$. The cofre ov measure theory is Lebesgue measure. Its most important generalization is the Haar measure for locally-compact topological and Lie groups.

$p(A \in \Delta; \psi) := \langle\psi|\, \widehat{E}[A \in \Delta]\, |\psi\rangle$. This means that upon measurement of the physical quantity A, one will find the measurement result to lie in Δ with probability $p(A \in \Delta; \psi)$. This *interpretation depends on measurements and an external observer.* Moreover, the measurement devices (and the observer) are described in terms of classical physics, not quantum physics [42]. We will elaborate on these ideas in the following sections. This will lead us to adopt a sophisticated realist interpretation that builds on the traditional instrumentalist view of quantum theory (under the Copenhagen interpretation) by also ascribing to the mathematical formalism genuine, though provisional, explanatory potential. In later sections we will apply this interpretation to *crowd behavior.*

This concludes our brief review of topos theory, which will be used in the following sections. For more technical details, see e.g. [24, 27, 28, 29].

4 Propositional Languages and Crowd Dynamics

4.1 Three Interpretations of Propositions

We start by considering the way in which *propositions* arise, and are manipulated, in crowd mechanics. Then, to each crowd system S there is associated a set of real-valued mechanical quantities, such as energy, momentum, position, angular momentum etc. The associated propositions are of the form '$A \in \Delta$', where A is a crowd quantity, and Δ is a Borel subset of $\mathbb{R}$.

From a conceptual perspective, the proposition '$A \in \Delta$' can be read in at least two, very different, ways [38]. To this account we also add a third view, which affords us with a sophisticated basis for studying propositions in crowd behaviour. The essential difference between these views is not so much Ontological (concerned with existence) as it is Epistemological (concerned with the nature of knowledge) and Methodological (concerned with the way in which knowledge is obtained). The difference really matters: whichever view we adopt fundamentally decides the meaning of our intuitionistic language for crowd behaviour:

(i) **The naive realist interpretation:** "The crowd quantity A *has* a value, and that value lies in Δ." We can evaluate the truth of our propositions absolutely.

(ii) **The instrumentalist interpretation:** "If a measurement is made of the crowd quantity A, the result will be found to lie in Δ." Our propositions are instruments.

(ii) **The sophisticated realist interpretation:** "We think the crowd has a quantity A, and its measurement will produce a value that lies in Δ." Our propositions have explanatory power, albeit provisional and limited.

The former is the familiar, 'commonsense' understanding of propositions in both crowd mechanics and daily life. This view holds that reality is pretty much how our widely accepted everyday experience – literally, our common sense – suggests it should be. This view asserts that there is a reality independent of our observing of it, that propositions about reality can be asserted to be true by observing that reality, that the existence and nature of the objects being observed is independent of the observation, and that as a result, we perceive the world through our senses pretty much as it is. Knowledge is therefore basically cumulative.

The second view denies that the truth content of our propositions about reality can be evaluated in an absolute sense, and sees those propositions instead as essentially an opaque computational (or, in other words, inference) system. In this view, our propositions constitute an engine that takes descriptions of specific circumstances and produces predictions of 'observations', which can be compared with actual observations of reality. In the instrumentalist view, we value our propositions by their usefulness, not by any notion of correctness. In other words, the value of our propositions is the collective value of the predictions they allow us to make. Instrumentalism underpins the Copenhagen interpretation of quantum theory.

The third view can be seen as a kind of optimistic development of the second, in which the denial that our propositions have truth value with respect to reality is relaxed to an energetic scepticism. This view inherits from instrumentalism the view that our systems of propositions are inference systems for producing predictions, yet here we treat the inference or computation in a more transparent manner; that is, we ascribe to our propositions provisional explanatory power.[40] The value of our propositions is the collective value of the explanation they afford, not just of the predictions they allow us to make.

It is this view that we will use later for analyzing crowd psycho-physical behavior. We will now study the role of propositions in crowd mechanics more carefully, particularly in the context of 'sophisticated realist' interpretations.

4.2 The Propositional Language $\mathcal{PL}(S)$

4.2.1 Intuitionistic Logic and the Definition of $\mathcal{PL}(S)$

We are going to construct a formal language, $\mathcal{PL}(S)$, with which to express propositions about a crowd system S and to make deductions concerning

---

[40] "... there are profound differences between pure theories and technological computation rules, and [that] instrumentalism can give a perfect description of these rules but is quite unable to account for the difference between them and the theories" [Karl Popper, Conjectures and Refutations]. Arguably, instrumentalism alone is sufficient to establish the basis for our intuitionistic language of crowd systems; however, we maintain that such a language should not just provide predictions, but it should also provide meaningful description of the phenomena. In other words, we hold that instrumentalism alone provides us with a view of the application and use of our crowd dynamics language that is just too mechanical. This contention is also consistent with criticisms of instrumentalism on the basis that it clashes with the methods of many scientists (importantly including physicists) [Bonner W.B, British Journal for the Philosophy of science 1958 VIII(32):291-294].

this system. Our intention is to define the meaning of these propositions in a 'sophisticated realist' way: we will be expressing provisional assertions about what we might observe and how and why we think we will observe those particular things. Compared with an instrumentalist perspective, we are intending not only that we are deriving predictions about observations we might make, but also that the formal derivation (i.e., proof-theoretic aspects of our formal logical system) provides useful – albeit provisional – supporting explanation of those predictions and thus of the workings of the crowd system. This is an endeavor whose mathematical underpinning lies in constructing a representation of $\mathcal{PL}(S)$ in a Heyting algebra $\mathfrak{H}$ that is part of the mathematical framework involved in the application of a particular theory-type to S. For further reading we suggest [67, 68, 69, 70, 71, 72].

The first step is to construct the set, $\mathcal{PL}(S)_0$, of all strings of the form '$A \in \Delta$' where A is a crowd quantity of the system S, and Δ is a Borel subset of the real line $\mathbb{R}$. Note that what has here been called a 'crowd quantity' could better (but more clumsily) be termed the 'name' of the crowd quantity. For example, when we talk about the 'energy' of a crowd system, the word 'energy' is the same, and functions in the same way in the formal language, irrespective of the details of the actual Hamiltonian energy function of the crowd system [38].

The strings '$A \in \Delta$' are taken to be the *primitive propositions* about the crowd system S and are used to define 'sentences'. More precisely, a new set of symbols $\{\neg, \wedge, \vee, \Rightarrow\}$ is added to the language, and then a *sentence* is defined *inductively* by the following rules [54]:

1. Each primitive proposition '$A \in \Delta$' in $\mathcal{PL}(S)_0$ is a sentence.
2. If α is a sentence, then so is $\neg\alpha$.
3. If α and β are sentences, then so are $\alpha \wedge \beta$, $\alpha \vee \beta$, and $\alpha \Rightarrow \beta$.

The collection of all sentences, $\mathcal{PL}(S)$, is an elementary formal language that can be used to express and manipulate propositions about the crowd system S. Note that the symbols $\neg$, $\wedge$, $\vee$, and $\Rightarrow$ have no explicit meaning, although the implicit intention is that they should stand for 'not', 'and', 'or' and 'implies', respectively. This implicit meaning will become explicit when a representation of $\mathcal{PL}(S)$ is constructed as part of the application of a theory-type to S. Note also that $\mathcal{PL}(S)$ is a *propositional* language only: it does not contain the quantifiers '$\forall$' or '$\exists$'. To include them requires a higher-order language.[41] We will return to this in our discussion of the local typed language $\mathcal{TL}(S)$.

The next step arises because $\mathcal{PL}(S)$ is not only a vehicle for expressing propositions about the crowd system S: we also want to *reason* with it about the system. To achieve this, a series of axioms for a deductive logic must be added to $\mathcal{PL}(S)$. This could be either classical logic or intuitionistic logic, but we select the latter since it allows a larger class of representations/models, including representations in topoi in which the law of excluded middle fails.

---

[41] Note that we are really using a natural deduction approach here, as opposed to an axiomatic approach.

Recall from Introduction that the axioms for intuitionistic logic consist of a finite collection of sentences in $\mathcal{PL}(S)$ (for example, $\alpha \wedge \beta \Rightarrow \beta \wedge \alpha$), plus a single rule of inference, *modus ponens* (the 'rule of detachment'), which says that from α and $\alpha \Rightarrow \beta$ the sentence β may be derived.

Others axioms might be added to $\mathcal{PL}(S)$ to reflect the implicit meaning of the primitive proposition '$A \in \varDelta$': i.e., "A has a value, and that value lies in $\varDelta \subseteq \mathbb{R}$." For example, the sentence '$A \in \varDelta_1 \wedge A \in \varDelta_2$' ('$A$ belongs to $\varDelta_1$' *and* 'A belongs to $\varDelta_2$') might seem to be equivalent to '$A \in \varDelta_1 \cap \varDelta_2$' ('$A$ belongs to $\varDelta_1 \cap \varDelta_2$'). A similar remark applies to '$A \in \varDelta_1 \vee A \in \varDelta_2$'.

Thus, along with the axioms of intuitionistic logic and modus ponens, we might be tempted to add the following axioms [38]:

$$A \in \varDelta_1 \wedge A \in \varDelta_2 \Leftrightarrow A \in \varDelta_1 \cap \varDelta_2, \tag{41}$$

$$A \in \varDelta_1 \vee A \in \varDelta_2 \Leftrightarrow A \in \varDelta_1 \cup \varDelta_2. \tag{42}$$

These axioms are consistent with the intuitionistic logical structure of $\mathcal{PL}(S)$. We will see later the extent to which the axioms (41–42) are compatible with the topos representations of crowd mechanics and of crowd behavior.

In classical logic, this proposition, '$\neg(A \in \varDelta)$',[42] is equivalent to 'A belongs to $\mathbb{R}\backslash\varDelta$', where $\mathbb{R}\backslash\varDelta$ denotes the set-theoretic complement of $\varDelta$ in $\mathbb{R}$. This suggests augmenting (41–42) with a third axiom:

$$\neg(A \in \varDelta) \Leftrightarrow A \in \mathbb{R}\backslash\varDelta. \tag{43}$$

However, applying '$\neg$' to both sides of (43) gives

$$\neg\neg(A \in \varDelta) \Leftrightarrow A \in \mathbb{R},$$

because of the set-theoretic result: $\mathbb{R}\backslash(\mathbb{R}\backslash\varDelta) = \varDelta$. But in an intuitionistic logic we do not have $\alpha \Leftrightarrow \neg\neg\alpha$ but only $\alpha \Rightarrow \neg\neg\alpha$, and so (43) could be false in a Heyting-algebra representation of $\mathcal{PL}(S)$ that was not Boolean. Therefore, adding (43) as an axiom in $\mathcal{PL}(S)$ is not indicated if representations are to be sought in non-Boolean topoi.

4.2.2 Representations of $\mathcal{PL}(S)$

To use the language $\mathcal{PL}(S)$ 'for real' it must be represented in the concrete mathematical structure that arises when a theory-type is applied to S. Such a representation π maps each of the primitive propositions α in $\mathcal{PL}(S)_0$ to an element $\pi(\alpha)$ of some (possibly Boolean) Heyting algebra $\mathfrak{H}$, whose specification is part of the theory [38]. In crowd mechanics, the propositions are represented in the Boolean algebra of all Borel subsets of the crowd phase space (see [18, 19, 20]).

---

[42] The parentheses ( ) are not symbols in the language; they are just a way of grouping letters and sentences.

The representation of the primitive propositions can be extended recursively to all of $\mathcal{PL}(S)$ with the aid of the following rules [54]:[43]

$$(a)\ \pi(\alpha \vee \beta) := \pi(\alpha) \vee \pi(\beta), \tag{44}$$

$$(b)\ \pi(\alpha \wedge \beta) := \pi(\alpha) \wedge \pi(\beta), \tag{45}$$

$$(c)\ \pi(\neg\alpha) := \neg\pi(\alpha), \tag{46}$$

$$(d)\ \pi(\alpha \Rightarrow \beta) := \pi(\alpha) \Rightarrow \pi(\beta). \tag{47}$$

This extension of π from $\mathcal{PL}(S)_0$ to $\mathcal{PL}(S)$ is consistent with the axioms for the intuitionistic, propositional logic of the language $\mathcal{PL}(S)$. More precisely, these axioms become tautologies: i.e., they are all represented by the maximum element 1 in the Heyting algebra. By construction, the map $\pi : \mathcal{PL}(S) \to \mathfrak{H}$ is then a representation of $\mathcal{PL}(S)$ in the Heyting algebra $\mathfrak{H}$. A logician would say that $\pi : \mathcal{PL}(S) \to \mathfrak{H}$ is an $\mathfrak{H}$-*valuation*, or $\mathfrak{H}$-*model*, of the language $\mathcal{PL}(S)$.

Note that different crowd systems S can have the same language. For example, consider a simple crowd agent of a point-particle type, moving in one dimension, with a Hamiltonian $H = \frac{p^2}{2m} + V(x)$. Different potentials $V(x)$ correspond to different crowd systems, but the mechanical quantities for these systems, or, more precisely, the 'names' of these quantities, for example, 'energy', 'position', 'momentum', are the same for them all. Consequently, the language $\mathcal{PL}(S)$ is independent of $V(x)$. However, the *representation* of, say, the proposition "$H \in \Delta$', with a specific subset of the state space *will* depend on the details of the Hamiltonian.

Clearly, a major consideration in using the language $\mathcal{PL}(S)$ is choosing the Heyting algebra in which the representation takes place. A fundamental result in topos theory is that the set of all sub-objects of any object in a topos is a Heyting algebra: these are the Heyting algebras with which we will be concerned [38].

Beyond the language $\mathcal{PL}(S)$ and its representation π, lies the question of whether or not a proposition is true. This requires the concept of a 'crowd state' which, when specified, yields 'truth values' for the primitive propositions in $\mathcal{PL}(S)$. These are then extended recursively to the rest of $\mathcal{PL}(S)$. In crowd mechanics, the possible truth values are just *true* or $false$.

There is also the question of 'how things change in time'. In the form presented above, the language $\mathcal{PL}(S)$ may seem geared towards a 'canonical' perspective in so far as the propositions concerned are, presumably, to be asserted at a particular moment of time, and, as such, deal with the values of physical quantities at that time. In other words, the underlying spatio-temporal perspective seems thoroughly 'Newtonian'. This is partly true; but only partly, since the phrase 'crowd quantity' can have meanings other than the canonical one. For example, one could talk about the 'time average of

---

[43] Note that, on the left hand side of (44–47), the symbols $\{\neg, \wedge, \vee, \Rightarrow\}$ are elements of the language $\mathcal{PL}(S)$, whereas on the right hand side they are the logical connectives in the Heyting algebra $\mathfrak{H}$ in which the representation takes place.

momentum', and call that a crowd quantity. In this case, the propositions would be about *histories* of the system, not just 'the way things are' at a particular moment in time.

We will return to these extended versions of the formalism in our discussion of the higher-order language, $\mathcal{TL}(S)$. However, for the moment let us focus on the canonical perspective, and the associated question of how time dependence is to be incorporated. This can be addressed in various ways.

One possibility is to attach a time label t to the crowd quantities, so that the primitive propositions become of the form '$A_t \in \Delta$'. In this case, the language itself becomes time-dependent, so that we should write $\mathcal{PL}(S)_t$. One might not like the idea of adding external labels in the language and, indeed, in our discussion of the higher-order language $\mathcal{TL}(S)$ we will strive to eliminate such things. However, in the present case, in so far as $\Delta \subseteq \mathbb{R}$ is already an 'external' (to the language) entity, there seems no particular objection to adding another one [38].

If we adopt this approach, the representation π will map '$A_t \in \Delta$' to a time-dependent element $\pi(A_t \in \Delta)$ of the Heyting algebra $\mathfrak{H}$.[44] However, this suggests another option, which is to keep the language time-independent, but allow the representation to be time-dependent. In that case, $\pi_t(A \in \Delta)$ will again be a time-dependent member of $\mathfrak{H}$.

Another approach is to let the 'truth object' in the theory be time-dependent: this corresponds to a type of quantum *Schrödinger picture.*

4.2.3 The Representation of $\mathcal{PL}(S)$ in Crowd Mechanics

Let us now look at the representation of $\mathcal{PL}(S)$ that corresponds to crowd mechanics. In this case, the topos involved is just the category **Set** of sets and functions between sets.

We will denote by π_{cl} the representation of $\mathcal{PL}(S)$ that describes the classical, Hamiltonian mechanics of a crowd system S, whose state-space is a symplectic (or Poisson) manifold $\mathcal{S}$ (see, e.g. [3]). We denote by $\breve{A} : \mathcal{S} \to \mathbb{R}$ the real-valued function[45] on $\mathcal{S}$ that represents the crowd quantity A. Then the representation π_{cl} maps the primitive proposition '$A \in \Delta$' to the subset of $\mathcal{S}$ given by [38]

$$\begin{aligned}\pi_{\mathrm{cl}}(A \in \Delta) &:= \{s \in \mathcal{S} \mid \breve{A}(s) \in \Delta\} \\ &= \breve{A}^{-1}(\Delta). \end{aligned} \tag{48}$$

This representation can be extended to all the sentences in $\mathcal{PL}(S)$ with the aid of (44–47). Note that, since Δ is a Borel subset of $\mathbb{R}$, $\breve{A}^{-1}(\Delta)$ is a Borel subset of the state-space $\mathcal{S}$. Hence, in this case, $\mathfrak{H}$ is equal to the Boolean algebra of all Borel subsets of $\mathcal{S}$.

---

[44] One could say that this is a type of quantum *Heisenberg picture.*

[45] In practice, $\breve{A}$ is required to be measurable, or smooth, depending on the type of crowd's physical quantity that A is.

We note that, for all Borel subsets $\varDelta_1, \varDelta_2$ of $\mathbb{R}$ we have [38]

$$\breve{A}^{-1}(\varDelta_1) \cap \breve{A}^{-1}(\varDelta_2) = \breve{A}^{-1}(\varDelta_1 \cap \varDelta_2), \tag{49}$$

$$\breve{A}^{-1}(\varDelta_1) \cup \breve{A}^{-1}(\varDelta_2) = \breve{A}^{-1}(\varDelta_1 \cup \varDelta_2), \tag{50}$$

$$\neg \breve{A}^{-1}(\varDelta_1) = \breve{A}^{-1}(\mathbb{R} \backslash \varDelta_1), \tag{51}$$

and hence all three conditions (41–43) that we discussed earlier can be added consistently to the language $\mathcal{PL}(S)$.

Consider now the assignment of truth values to the propositions in this theory. This involves the idea of a 'state' which, in crowd mechanics, is simply an element s of the crowd state space $\mathcal{S}$. Each state s assigns to each primitive proposition '$A \in \varDelta$', a truth value $\nu\big(A \in \varDelta; s\big)$, which lies in the set $\{false, true\} := \{\bot, \top\}$, where $\bot$ is false in all interpretations, $\top$ is true in all interpretations and 'interpretation' here specifically means an assignment of semantic meaning to one or more propositions (not to be confused with the future use of the word 'interpretation').

4.2.4 The Failure to Represent $\mathcal{PL}(S)$ in Crowd Behavior

The procedure above that works so easily for crowd mechanics fails completely if one tries to apply it to crowd behavior.

In quantum-like crowd behavior, a mechanical crowd quantity A is represented by a self-adjoint operator $\hat{A}$ on a crowd Hilbert space $\mathcal{H}$, and the proposition '$A \in \varDelta$' is represented by the projection operator $\hat{E}[A \in \varDelta]$ which projects onto the subset $\varDelta$ of the spectrum of $\hat{A}$ i.e.,

$$\pi(A \in \varDelta) := \hat{E}[A \in \varDelta]. \tag{52}$$

The set of all projection operators $\mathcal{P}(\mathcal{H})$ in $\mathcal{H}$ has a 'logic' of its own, the 'quantum logic'[46] of the Hilbert space $\mathcal{H}$, but this is incompatible with the intuitionistic logic of the language $\mathcal{PL}(S)$, and the representation (52).

Indeed, since the 'logic' $\mathcal{P}(\mathcal{H})$ is non-distributive, there will exist non-commuting operators $\hat{A}, \hat{B}, \hat{C}$, and Borel subsets $\varDelta_A, \varDelta_B, \varDelta_C$ of $\mathbb{R}$ such that [38]

$$\hat{E}[A \in \varDelta_A] \wedge \Big(\hat{E}[B \in \varDelta_B] \vee \hat{E}[C \in \varDelta_C]\Big) \neq \\ \Big(\hat{E}[A \in \varDelta_A] \wedge \hat{E}[B \in \varDelta_B]\Big) \vee \Big(\hat{E}[A \in \varDelta_A] \wedge \hat{E}[C \in \varDelta_C]\Big)$$

while, on the other hand, the logical equivalence (or, bi-implication)

$$\alpha \wedge (\beta \vee \gamma) \Leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$$

can be deduced from the axioms of the language $\mathcal{PL}(S)$.

---

[46] For an excellent survey of quantum logic see [56]. This includes a discussion of a first-order axiomatization of quantum logic, and with an associated sequent calculus.

This failure of distributivity bars any naive realist interpretation of quantum logic. If an instrumentalist interpretation is used instead, the spectral projectors $\hat{E}[A \in \Delta]$ now represent propositions about what would happen *if* a measurement is made, not propositions about what is 'actually the case'. And when a state is specified, this does not yield actual truth values but only the Born-rule probabilities of getting certain results [38].

4.3 Instrumentalism, Tableau and Resolution

4.3.1 Instrumentalism

Recall that the instrumentalist position views our propositions as constituting a kind of computational machinery for generating predictions about what we might observe, given sufficient specification of the circumstances of our hypothetical observation. Our sophisticated realist position inherits this feature while also contending that our language has explanatory power: it conjectures not just the "what" but also the "how" and the "why". Intuitionistic logic is interesting in this setting: unlike a classical proof, an intuitionistic proof constitutes a computable function (or, if you prefer, an effective procedure).[47] In other words, intuitionistic logics allow us, at least in principle, to synthesize computable functions.

In this section, we briefly discuss the kind of methods that we might use to implement proof infrastructure to concretely support the position that our formal statements - propositions - about crowd dynamics constitute computational mechanisms for generating predictions. We are interested in two closely related modes of operation. Firstly, we may want to be able to prove (or disprove) a statement or set of statements expressed in our language. Secondly, we may want to generate models, which are possible interpretations, or assignments of semantic meaning that make a statement or statements true in our language[48]. Our goal here is not to provide a complete overview of theorem proving and model generation, nor to offer a particular method or even set of

---

[47] We are encouraging a functional notion of computation here, such as might be expressed in a typed lambda calculus, rather than a state-oriented view such as a Turing Machine. The functional view is much more abstract, and aligns closely with other aspects of our approach. For instance, Type Theory, which can be viewed essentially as an application of Category Theory, arguably sits much more comfortably with the functional view of computation, as demonstrated by the vast difference in expressive power of type systems in functional languages to that of state-oriented languages (including object-oriented languages).

[48] Note that this does not imply that those statements are true of crowd phenomenon, only that they are consistent with our understanding of it as expressed in our intuitionistic language of crowd dynamics. Whether and to what degree the statements correspond with our crowd phenonomena is an empirical matter, not a matter of formal logic. This is another example of the careful distinction our sophisticated realist view makes between the phenomenon in the real world and our theories that attempt to explain it.

methods, but rather to give a flavour both of the difficulties involved and of the two broad types of approaches discussed in the literature. The reason for this is that there is really no universally satisfactory technique, and the details of what works best vary intricately according to the specifics of the intuitionistic logic and the kinds of statements we might be interested in.

Proof and model generation in intuitionistic logic is much more difficult than in classical logics. In intuitionistic logics, we cannot in general blindly use Herbrand functions (often also called "skolemization") except in some special cases, and because of the asymmetries and relative weakness of intuitionistic languages,[49] we have no convenient normal forms such as conjunctive normal form, disjunctive normal form, negation normal form or prenex normal form. There are relatively few classes known for intuitionistic logics that are both interesting and decidable.[50]

Type theories are fundamentally higher-order, yet there are useful first-order fragments; proof search in such fragments amounts to proof search in first-order intuitionistic logic. While we cannot expect that even most complicated though provable statements in intuitionistic logic will be amenable to automatic methods, we do expect that theorem proving and model generating methods could provide substantial assistance for deducing predictions and explanations in our language of crowd dynamics. Our case is also helped by the significantly softer resource demands of our problem: nobody is happy to wait for a couple of hours for a compilation of a modest program, but we would likely be cope well enough if solving a substantial system of statements about crowd behaviour for its models took the whole weekend.

We continue using Gentzen-type natural deduction systems, or sequent calculi. In such systems, the axioms and theorems are not single statements in the underlying logic, but pairs of statements that form inference rules, called "sequents". In Gentzen-type systems, deduction is is oriented around provability rather than around truth judgements. The sequent

$$\varGamma \vdash \varDelta$$

means that if we accept the antecedent $\varGamma$ then we may conclude the succedent $\varDelta$. Note that in general $\varGamma$ and $\varDelta$ range over multisets of formulas rather than sets of formulas. The proof system is defined by a set of inference rules of the form

$$\frac{P_1 P_2 \dots P_n}{C} (L)$$

---

[49] Note that we intend "weakness" and "strength" of logics in the manner standard in logic, which is to say in terms of validity of formulas in different logics: many tautologies (formulas that evaluate to "true" for all interpretations) of classical logics are not tautologies of corresponding intuitionistic logics.

[50] Read "decidable" as meaning that there exists a computational procedure that solves the problem of proving/disproving theorems or generating all models in all specific cases in the class.

where the sequents $P_1, P_2, \ldots, P_n$ are the premises, the sequent C is the consequence or conclusion, and L is a label naming the rule. The statement produced by the inference rule is usually called "the main formula", the parts of the main formula appearing in the premises are "side formulas", and other formulas in the rule are called "parametric formulas". For example, the rule

$$\frac{\Gamma \vdash A \; B, \Sigma \vdash D}{(A \Rightarrow B, \Gamma, \Sigma) \vdash D}(\Rightarrow\vdash)$$

basically states that we may conclude that D is provable given $A \Rightarrow B$, Γ, and Σ whenever we have that A is provable given Γ and that D is provable given B and Σ. Here, $A \Rightarrow B$ is the main formula, A and B are side formulas, and the remaining formulas Γ, Σ, D are parametric formulas. Most rules satisfy the "subformula property", in which every formula in the sequents of the premises will be a subformula or substitution instance of the formulas in the conclusion sequent. We will also assume the "eigenvariable condition" on rules involving a form of substitution for quantified variables, which essentially precludes substitutions that would cause capture of free variables by variable binding operators such as existential $\exists$ and universal quantifiers $\forall$. If these property holds for every rule of the system, then we have a "cut-free sequent calculus": this is important because the order in which the rules occur in a cut-free sequent calculus can be changed without changing the conclusion of the proof. We will consider only cut-free calculi here: both classical first-order logic and its intuitionistic counterpart enjoy the "cut-elimination property", which says that rules violating the subformula property can be eliminated.

4.3.2 Analytic Tableau

Analytic tableau[51] are a class of flexible computational procedures that are amenable to both proof search and model generation. Though historically tableau methods were developed in the context of the semantics of logical operators, they can also be seen as deriving from the cut-elimination theorem of proof theory. The search begins with the negative of the formula to be proved and branches backwards (that is, the sequent calculus rules are applied in a bottom-up fashion). In effect, analytic tableau work by attempting to refute. The search concludes when no further rules can be applied to any of the branches. Separate branches are semantically in disjunction, while nodes on the same branch are semantically in conjunction. Branches are said to be closed when they contain a contradiction, saturated when they are not closed but no further rules can be applied, and open otherwise.

The subformula property is important because it guarantees that there will only be a finite number of branches generated by the application of any particular rule. It is also useful in allowing us to apply rules in any order we like, and therefore we can apply rules that produce only one sub-branch first,

---

[51] We aim here to outline basic tableau methods for propositional and first-order intuitionistic logic.

effectively pushing rules that cause the tableau to fan out and the search procedure to explode exponentially towards the bottom of the tableau where their effect on the practical efficiency of the process is reduced. Because analytic tableau effectively decompose our starting statement into what semantically comprises a disjunction of conjunctions of literals, the approach is readily adaptable to model generation: effectively each conjunction of literals appearing at the bottom of a saturated branch is a model.

4.3.3 Resolution

In contrast to analytic tableau methods, resolution proof search methods (or "inverse methods") apply inference rules in a top-down direction.[52] Search begins with the set of axioms and we apply the inference rules are until we eventually derive the statement we wish to prove. Given that we can assume we have a cut-free sequent calculus, the proof search will only ever derive sequents that contain sub-formulas of the formula we wish to prove. The method is obviously complete: if there is a derivation of the formula we wish to prove from our axioms using the calculus, the forward-chaining application of all the possible rules in turn will in principle eventually find a derivation. However, as it stands this would be utterly useless in practice because of the combinatorial explosion in the derivation tree, so resolution relies on using strategies to limit the application of inference rules.

5 A Higher-Order, Typed Language for Crowd Behavior

5.1 The Basics of the Language $\mathcal{TL}(S)$

Now we want to consider the possibility of representing the physical quantities of a generic crowd system by morphisms in a topos other than **Set**.

The physical meaning of such a quantity is not clear, *a priori*. In such a situation it is no longer correct to work with a fixed value-space $\mathbb{R}$. Rather, the target-object, $\mathcal{R}_S$, is potentially topos-dependent, and therefore part of the 'representation'.

A powerful technique for allowing the quantity-value object to be system-dependent is to add a symbol '$\mathcal{R}$' to the language. Developing this line of thinking suggests that 'Σ', too, should be added, as should a symbol '$A : \Sigma \to \mathcal{R}$', to be construed as 'what it is' that is represented by the morphism

---

[52] We aim here to outline basic resolution methods for propositional and first-order intuitionistic logic.

We could possibly drop resolution and focus on analytic tableau alone, merely mentioning the existence of resolution methods and justifying the focus on analytic tableau because of their relative flexibility, good performance at relatively lower design and implementation complexity, and direct applicability to model generation as well as proof search.

in a topos. Similarly, there should be a symbol 'Ω', to act as the linguistic precursor to the sub-object classifier in the topos; in the topos **Set**, this is just the set $\{0, 1\}$.

The clean way of doing all this is to construct, what Bell [28] calls, a 'local language'. Our basic assumption is that a unique local language, $\mathcal{TL}(S)$, is associated with each system S. Physical theories of S then correspond to representations of $\mathcal{TL}(S)$ in appropriate topoi.

We first consider the minimal set of symbols needed to handle elementary crowd dynamics. The symbols for the local language $\mathcal{TL}(S)$ are defined recursively as follows [38]:

1. (a) The basic *type symbols* are $1, \Omega, \Sigma, \mathcal{R}$. The last two, Σ and $\mathcal{R}$, are known as *ground-type symbols.* They are the linguistic precursors of the state object, and quantity-value object, respectively.
   If $T_1, T_2, \ldots, T_n$, $n \geq 1$, are type symbols, then so is[53] $T_1 \times T_2 \times \cdots \times T_n$.
   (b) If T is a type symbol, then so is PT.
   (c) For each type symbol, T, there is associated a countable set of *variables of type* T.
   (d) There is a special symbol $*$.
2. (a) To each pair (T_1, T_2) of type symbols there is associated a set, $F_{\mathcal{TL}(S)}(T_1, T_2)$, of *function symbols.* Such a symbol, A, is said to have *signature* $T_1 \to T_2$; this is indicated by writing $A : T_1 \to T_2$.
   (b) Some of these sets of function symbols may be empty. However, particular importance is attached to the set, $F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$, of function symbols $A : \Sigma \to \mathcal{R}$, and we assume this set is non-empty.

The function symbols $A : \Sigma \to \mathcal{R}$ represent the 'physical quantities' of the system, and hence $F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$ will depend on the system. In fact, the only parts of the language that are system-dependent are these function symbols.

For example, if S_1 is an agent moving in one dimension, the set of physical quantities could be chosen to be $F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) = \{x, p, H\}$ which represent the position, momentum, and energy of the agent system. On the other hand, if S_2 is an agent moving in three dimensions, we could have $F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) = \{x, y, z, p_x, p_y, p_z, H\}$ to allow for three-dimensional position and momentum of the agent. Or, we could decide to add angular momentum too, to give the set $F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) = \{x, y, z, p_x, p_y, p_z, J_x, J_y, J_z, H\}$.[54]

It should be emphasized that this list of symbols is minimal and one may want to add more. One obvious, general, example is a type symbol $\mathbb{N}$, to be interpreted as the linguistic analogue of the natural numbers. The language could then be augmented with the axioms of *Peano arithmetic.*

---

[53] By definition, if $n = 0$ then $T_1 \times T_2 \times \cdots \times T_n := 1$.

[54] As with the propositional language $\mathcal{PL}(S)$, the fact that a given system has a specific Hamiltonian—expressed as a particular function of position and momentum coordinates—is not something that is to be coded into the language: instead, such system dependence arises in the choice of *representation* of the language. This means that many different systems can have the same local language.

The next step is to enumerate the 'terms' in the language, together with their associated types [28, 30]:

1. (a) For each type symbol T, the variables of type T are terms of type T.
   (b) The symbol $*$ is a term of type 1.
   (c) A term of type Ω is called a *formula*; a formula with no free variables is called a *sentence.*
   (d) Closure: nothing else is a formula.
2. If A is function symbol with signature $T_1 \to T_2$, and t is a term of type T_1, then $A(t)$ is term of type T_2.
   In particular, if $A : \Sigma \to \mathcal{R}$ is a physical quantity, and t is a term of type Σ, then $A(t)$ is a term of type $\mathcal{R}$.
3. (a) If $t_1, t_2, \ldots, t_n$ are terms of type $T_1, T_2, \ldots, T_n$, then $\langle t_1, t_2, \ldots, t_n \rangle$ is a term of type $T_1 \times T_2 \times \cdots \times T_n$.
   (b) If t is a term of type $T_1 \times T_2 \times \cdots \times T_n$, and if $1 \leq i \leq n$, then $(t)_i$ is a term of type T_i.
   (c) If ω is a term of type Ω, and $\tilde{x}$ is a variable of type T, then $\{\tilde{x} \mid \omega\}$ is a term of type PT.
   (d) If t_1, t_2 are terms of the same type, then $t_1 = t_2$ is a term of type Ω.
   (e) If t_1, t_2 are terms of type T, PT respectively, then $t_1 \in t_2$ is a term of type Ω.

Note that the logical operations are not included in the set of symbols. Instead, they can all be defined using what is already given. For example, (i) *true* $:= (* = *)$; and (ii) if α and β are terms of type Ω, then[55]

$$\alpha \wedge \beta := \big(\langle \alpha, \beta \rangle = \langle \mathrm{true}, \mathrm{true} \rangle\big).$$

Thus, in terms of the original set of symbols, we have

$$\alpha \wedge \beta := \big(\langle \alpha, \beta \rangle = \langle * = *, * = * \rangle\big), \quad \text{etc.}$$

Let A be a physical quantity in the set $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$, and therefore a function symbol of signature $\Sigma \to \mathcal{R}$. In addition, let $\tilde{\Delta}$ be a variable (and therefore a term) of type $P\mathcal{R}$; and let $\tilde{s}$ be a variable (and therefore a term) of type Σ. Then some terms of particular interest to us are the following [38]:

1. $A(\tilde{s})$ is a term of type $\mathcal{R}$ with a free variable, $\tilde{s}$, of type Σ.
2. '$A(\tilde{s}) \in \tilde{\Delta}$' is a term of type Ω with free variables (i) $\tilde{s}$ of type Σ; and (ii) $\tilde{\Delta}$ of type $P\mathcal{R}$.
3. $\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}$ is a term of type $P\Sigma$ with a free variable $\tilde{\Delta}$ of type $P\mathcal{R}$.

As we will see, $\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}$ and '$A(\tilde{s}) \in \tilde{\Delta}$' are (closely related) analogues of the primitive propositions '$A \in \Delta$' in the propositional language $\mathcal{PL}(S)$. However, there is a crucial difference. In $\mathcal{PL}(S)$, the 'Δ' in '$A \in \Delta$' is a

---

[55] The parentheses ( ) are not symbols in the language, they are just a way of grouping letters and sentences.

specific subset of the external (to the language) real line $\mathbb{R}$. On the other hand, in the local language $\mathcal{TL}(S)$, the '$\tilde{\Delta}$' in '$A(\tilde{s}) \in \tilde{\Delta}$' is an *internal* variable within the language.

Finally, to make the language $\mathcal{TL}(S)$ into a *deductive system* we need to add a set of appropriate *axioms* and *inference rules.* The former are expressed using *sequents*:[56] defined as expressions of the form $\Gamma : \alpha$ where α is a formula (a term of type Ω) and Γ is a set of such formula. The intention is that '$\Gamma : \alpha$' is to be read intuitively as "the collection of formula in Γ 'imply' α". If Γ is empty we just write $: \alpha$.

The basic axioms include things like '$\alpha : \alpha$' (tautology), and '$: \tilde{t} \in \{\tilde{t} \mid \alpha\} \Leftrightarrow \alpha$' (comprehension) where $\tilde{t}$ is a variable of type T. These axioms[57] and the rules of inference (sophisticated analogues of *modus ponens*) give rise to a deductive system using intuitionistic logic. For the details see [28, 30].

However, for applications in crowd dynamics we could add extra axioms (in the form of sequents). For example, perhaps the quantity-value object should always be an Abelian-group object[58]? This can be coded into the language by adding the axioms for an Abelian group structure for $\mathcal{R}$. This involves the following steps [38]:

1. Add the following symbols:
    (a) A 'unit' function symbol $0 : 1 \to \mathcal{R}$; this will be the linguistic analogue of the unit element in an Abelian group.
    (b) An 'addition' function symbol $+ : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$.
    (c) An 'inverse' function symbol $- : \mathcal{R} \to \mathcal{R}$
2. Then add axioms like '$: \forall \tilde{r}\big(+\langle \tilde{r}, 0(*)\rangle = \tilde{r}\big)$' where $\tilde{r}$ is a variable of type $\mathcal{R}$, and so on.

---

[56] We are choosing this approach here; alternatives include axiomatic approache4s using *axiom schemes.*

[57] The complete set of axioms is [28]:

Tautology: $\alpha = \alpha$

Unity : $\tilde{x}_1 = *$ where $\tilde{x}_1$ is a variable of type 1.

Equality: $x = y, \alpha(\tilde{z}/x) : \alpha(\tilde{z}/y)$. Here, $\alpha(\tilde{z}/x)$ is the term α with $\tilde{z}$ replaced by the term x for each free occurrence of the variable $\tilde{z}$. The terms x and y must be of the same type as $\tilde{z}$.

Products: $: (\langle x_1, \ldots, x_n\rangle)_i = x_i$

$: x = \langle (x)_1, \ldots, (x)_n\rangle$

Comprehension: $: \tilde{t} \in \{\tilde{t} \mid \alpha\} \Leftrightarrow \alpha.$

[58] One could go even further and add the axioms for real numbers. In this case, in a representation of the language in a topos τ, the symbol $\mathcal{R}$ is mapped to the real-number object in the topos (if there is one). However, the example of quantum theory suggests that this is inappropriate [40].

For another example, consider an agent moving in three dimensions, with the function symbols $F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) = \{x, y, z, p_x, p_y, p_z, J_x, J_y, J_z, H\}$. As $\mathcal{TL}(S)$ stands, there is no way to specify, for example, that '$J_x = yp_z - zp_y$'. Such relations can only be implemented in a *representation* of the language. However, if this relation is felt to be 'universal' (i.e., it holds in all physically-relevant representations) then it could be added to the language with the use of extra axioms.

One of the delicate decisions that has to be made about $\mathcal{TL}(S)$ is what extra axioms to add to the base language. Too few, and the language lacks content; too many, and representations of potential physical significance are excluded. This is one of the places in the formalism where a degree of physical insight is necessary!

5.2 Representing $\mathcal{TL}(S)$ in a Topos

The construction of a theory of the system S involves choosing a representation/model, ϕ, of the language $\mathcal{TL}(S)$ in a topos τ_ϕ. The choice of both topos and representation depend on the theory-type being used.

For example, consider a crowd subsystem S that can be treated using both crowd mechanics and crowd behavior, such as an agent moving in three dimensions. Then, for the application of the theory-type 'crowd mechanics', in a representation denoted σ, the topos τ_σ is **Set**, and Σ is represented by the symplectic manifold $\Sigma_\sigma := T^*\mathbb{R}^3$.

On the other hand, for the application of the theory-type 'crowd behavior', τ_ϕ is the topos, $\mathbf{Set}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$, of presheaves over the category $\mathcal{V}(\mathcal{H})$, where $\mathcal{H} \simeq L^2(\mathbb{R}^3, d^3x)$ is the Hilbert space of the system S. In this case, Σ is represented by $\Sigma_\phi := \underline{\Sigma}$, where $\underline{\Sigma}$ is the spectral presheaf; this representation is discussed at length in [39, 40]. For both theory types, the *details* of, for example, the Hamiltonian, are coded in the representation.

We now list the τ_ϕ-representation of the most significant symbols and terms in our language, $\mathcal{TL}(S)$ [28, 30].

1. (a) The ground type symbols Σ and $\mathcal{R}$ are represented by objects Σ_ϕ and $\mathcal{R}_\phi$ in τ_ϕ. These are identified physically as the state object, and quantity-value object, respectively.
   (b) The symbol Ω, is represented by $\Omega_\phi := \Omega_{\tau_\phi}$, the sub-object classifier of the topos τ_ϕ.
   (c) The symbol 1, is represented by $1_\phi := 1_{\tau_\phi}$, the terminal object in τ_ϕ.
2. For each type symbol PT, we have $(PT)_\phi := PT_\phi$, the power object of the object T_ϕ in τ_ϕ.
   In particular, $(P\Sigma)_\phi = P\Sigma_\phi$ and $(P\mathcal{R})_\phi = P\mathcal{R}_\phi$.
3. Each function symbol $A : \Sigma \to \mathcal{R}$ in $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$ (i.e., each physical quantity) is represented by a morphism $A_\phi : \Sigma_\phi \to \mathcal{R}_\phi$ in τ_ϕ.
   We will generally require the representation to be *faithful*: i.e., the map $A \mapsto A_\phi$ is one-to-one and onto, i.e., bijective.

4. A term of type Ω of the form '$A(\tilde{s}) \in \tilde{\Delta}$' (which has free variables $\tilde{s}, \tilde{\Delta}$ of type Σ and $P\mathcal{R}$ respectively) is represented by a morphism $[\![A(\tilde{s}) \in \tilde{\Delta}]\!]_\phi : \Sigma_\phi \times P\mathcal{R}_\phi \to \Omega_{\tau_\phi}$. In detail, this morphism is

   $$[\![A(\tilde{s}) \in \tilde{\Delta}]\!]_\phi = e_{\mathcal{R}_\phi} \circ \langle [\![A(\tilde{s})]\!]_\phi, [\![\tilde{\Delta}]\!]_\phi \rangle,$$

   where $e_{\mathcal{R}_\phi} : \mathcal{R}_\phi \times P\mathcal{R}_\phi \to \Omega_{\tau_\phi}$ is the usual evaluation map; $[\![A(\tilde{s})]\!]_\phi : \Sigma_\phi \to \mathcal{R}_\phi$ is the morphism A_ϕ; and $[\![\tilde{\Delta}]\!]_\phi : P\mathcal{R}_\phi \to P\mathcal{R}_\phi$ is the identity. Thus $[\![A(\tilde{s}) \in \tilde{\Delta}]\!]_\phi$ is the chain of morphisms:

   $$\Sigma_\phi \times P\mathcal{R}_\phi \xrightarrow{A_\phi \times id} \mathcal{R}_\phi \times P\mathcal{R}_\phi \xrightarrow{e_{\mathcal{R}_\phi}} \Omega_{\tau_\phi}. \tag{53}$$

   We see that the analogue of the 'Δ' used in the $\mathcal{PL}(S)$-propositions '$A \in \Delta$' is played by sub-objects of $\mathcal{R}_\phi$ (i.e., global elements of $P\mathcal{R}_\phi$) in the domain of the morphism in (53). These objects are, of course, representation-dependent (i.e., they depend on ϕ).
5. A term of type $P\Sigma$ of the form $\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}$ (which has a free variable $\tilde{\Delta}$ of type $P\mathcal{R}$) is represented by a morphism $[\![\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}]\!]_\phi : P\mathcal{R}_\phi \to P\Sigma_\phi$. This morphism is the power transpose[59] of $[\![A(\tilde{s}) \in \tilde{\Delta}]\!]_\phi$:

   $$[\![\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}]\!]_\phi = \ulcorner [\![A(\tilde{s}) \in \tilde{\Delta}]\!]_\phi \urcorner \tag{54}$$

6. A term, ω, of type Ω with no free variables is represented by a global element $[\![\omega]\!]_\phi : 1_{\tau_\phi} \to \Omega_{\tau_\phi}$. These will typically act as 'truth values' for propositions about the system.
7. Any axioms that have been added to the language are required to be represented by the morphism *true* $: 1_{\tau_\phi} \to \Omega_{\tau_\phi}$.

We emphasize that the decision to focus on the particular type of language that we have, is not an arbitrary one. Indeed, there is a deep connection between such languages and topos theory. In this context, we first note that to any local language, $\mathcal{L}$, there is associated a 'local set theory'. This involves defining an '$\mathcal{L}$-set' to be a term X of power type (so that expressions of the form $x \in X$ are meaningful) and with no free variables. Analogues of all the usual set operations can be defined on $\mathcal{L}$-sets. For example, if X, Y are $\mathcal{L}$-sets of type PT, one can define $X \cap Y := \{\tilde{x} \mid \tilde{x} \in X \wedge \tilde{x} \in Y\}$ where $\tilde{x}$ is a variable of type T.

Furthermore, each local set theory, $\mathcal{L}$, gives rise to an associated topos, $\mathcal{C}(\mathcal{L})$, whose objects are equivalence classes of $\mathcal{L}$-sets, where $X \equiv Y$ is defined to mean that the equation $X = Y$ (i.e., a term of type Ω with no free variables) can be proved using the sequent calculus of the language with its axioms. From this perspective, a representation of $\mathcal{TL}(S)$ in a topos τ is equivalent to a *functor* from the topos $\mathcal{C}(\mathcal{TL}(S))$ to τ.

---

[59] One of the basic properties of a topos is that there is a one-to-one correspondence between morphisms $f : A \times B \to \Omega$ and morphisms $\ulcorner f \urcorner : A \to PB := \Omega^B$. In general, $\ulcorner f \urcorner$ is called the *power transpose* of f. If $A \simeq 1$ then $\ulcorner f \urcorner$ is known as the *name* of the morphism $f : B \to \Omega$; see (17).

Conversely, for each topos τ there is a local language, $\mathcal{L}(\tau)$, whose ground-type symbols are the objects of τ, and whose function symbols are the morphisms in τ. It then follows that a representation of a local language, $\mathcal{L}$, in τ is equivalent to a 'translation' of $\mathcal{L}$ in $\mathcal{L}(\tau)$ [38].

Thus, a rather elegant way of summarizing what is involved in constructing a theory of physics is that we are *translating* the language, $\mathcal{T}\mathcal{L}(S)$, of the system in another local language, $\mathcal{T}\mathcal{L}(\tau)$. As we will see in the last section, the idea of translating one local language into another plays a central role in the discussion of composite systems and sub-systems [41].

5.3 Crowd Mechanics in the Local Language $\mathcal{T}\mathcal{L}(S)$

The quantum theory representation of $\mathcal{T}\mathcal{L}(S)$ was studied in [39, 40]. Here we will look at the concrete form of the expressions in the previous Section for the example of crowd mechanics. In this case, for all systems S, and all classical representations, σ, the topos τ_σ is **Set**. This representation of $\mathcal{T}\mathcal{L}(S)$ has the following ingredients:

1. (a) The ground-type symbol Σ is represented by a symplectic manifold Σ_σ that is the state-space for the system S.
   (b) The ground-type symbol $\mathcal{R}$ is represented by the real line, i.e., $\mathcal{R}_\sigma := \mathbb{R}$.
   (c) The type symbol $P\Sigma$ is represented by the set $P\Sigma_s$ of all subsets of the state space Σ_s
   The type symbol $P\mathcal{R}$ is represented by the set $P\mathbb{R}$ of all subsets of $\mathbb{R}$.
   (d) The type symbol Ω, is represented by $\Omega_{\mathbf{Set}} := \{0, 1\}$: the sub-object classifier in **Set**.
   (e) The type symbol 1, is represented by the singleton set, i.e., $1_{\mathbf{Set}} = \{*\}$: the terminal object in **Set**.
2. Each function symbol $A : \Sigma \to \mathcal{R}$, and hence each physical quantity, is represented by a real-valued function, $A_\sigma : \Sigma_\sigma \to \mathbb{R}$, on the state space Σ_σ.
3. The term '$A(\tilde{s}) \in \tilde{\Delta}$' of type Ω (where $\tilde{s}$ and $\tilde{\Delta}$ are free variables of type Σ and $P\mathcal{R}$ respectively) is represented by the function $[\![\, A(\tilde{s}) \in \tilde{\Delta} \,]\!]_s : \Sigma_s \times P\mathbb{R} \to \{0, 1\}$ that is defined by (c.f. (53))
$$[\![\, A(\tilde{s}) \in \tilde{\Delta} \,]\!]_s(s, \Delta) = \begin{cases} 1, & \text{if } A_\sigma(s) \in \Delta; \\ 0, & \text{otherwise.} \end{cases} \qquad \text{for all } (s, \Delta) \in \Sigma_s \times P\mathbb{R}. \tag{55}$$
4. The term $\{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\}$ of type $P\Sigma$ (where $\tilde{\Delta}$ is a free variable of type $P\mathcal{R}$) is represented by the function $[\![\, \{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\} \,]\!]_s : P\mathbb{R} \to P\Sigma_s$ that is defined by
$$\begin{aligned} [\![\, \{\tilde{s} \mid A(\tilde{s}) \in \tilde{\Delta}\} \,]\!]_s(\Delta) &:= \{s \in \Sigma_\phi \mid A_s(s) \in \Delta\} \\ &= A_s^{-1}(\Delta), \qquad \text{for all } \Delta \in P\mathbb{R}. \end{aligned}$$

5.4 Adapting the Language $\mathcal{TL}(S)$ to Other Types of Crowd System

Following [38, 39, 40, 41], our basic assumptions are:

1. Each crowd system, S, can be equipped with a local language, $\mathcal{TL}(S)$; and
2. Constructing an explicit theory of S in a particular theory-type is equivalent to finding a representation of $\mathcal{TL}(S)$ in a topos which may well be other than the topos of sets.

There are many situations in which the language is independent of the theory-type, and then, for a given system S, the different topos representations of $\mathcal{TL}(S)$, correspond to the application of the different theory-types to the same system S. We gave an example earlier of an agent moving in three dimensions: the crowd mechanics representation is in the topos **Set**; and, as shown in [40, 41], the quantum theory representation is in the presheaf topos $\mathbf{Set}^{\mathcal{V}(L^2(\mathbb{R}^3,\, d^3x))}$.

However, there are other situations where the relationship between the language and its representations is more complicated than this. In particular, there is the critical question about what features of the theory should go into the language, and what into the representation. Adding new features would begin by adding to, or changing, the set of ground-type symbols which generally represent the entities that are going to be of generic interest (such as a state object or quantity-value object). In doing this, extra axioms may also be introduced to encode the properties that the new objects are expected to possess in all the representations that are of physical interest [38].

For example, suppose we want to use our formalism to discuss space-time physics: where does the information about the space-time go? If the subject is classical field theory in a curved space-time, then the topos τ is **Set**, and the space-time manifold is part of the *background* structure. This makes it natural to have the manifold assumed in the representation; i.e., the information about the space-time is in the representation.

However, alternatively one can add a new ground type symbol, 'M', to the language, to serve as the linguistic progenitor of 'space-time'; thus M would have the same theoretical status as the symbols Σ and $\mathcal{R}$. A function symbol $\psi : M \to \mathcal{R}$ is then the progenitor of a physical field. In a representation ϕ, the object M_ϕ plays the role of 'space-time' in the topos τ_ϕ, and $\psi_\phi : M_\phi \to \mathcal{R}_\phi$ is the representation of a field in this theory.

Clearly, the language $\mathcal{TL}(S)$ says nothing about what sort of entity M_ϕ is, except in so far as such information is encoded in extra axioms. For example, if the subject is classical field theory, then $\tau_\phi = \mathbf{Set}$, and M_ϕ would be a standard differentiable manifold. On the other hand, if the topos τ_ϕ admits 'infinitesimals', then M_ϕ could be a manifold according to the language of *synthetic differential geometry* [55].

A fortiori, the same type of argument applies to the status of 'time' in a canonical theory. In particular, it is possible to add a ground type symbol,

$\mathcal{T}$, so that, in any representation, ϕ, the object $\mathcal{T}_\phi$ in the topos τ_ϕ is the analogue of the 'time-line' for that theory. For standard physics in **Set** we have $\mathcal{T}_\phi = \mathbb{R}$, but the form of $\mathcal{T}_\phi$ in a more general topos, τ_ϕ, would be a rich subject for speculation.

The addition of a 'time-type' symbol, $\mathcal{T}$, to the language $\mathcal{TL}(S)$ is a prime example of a situation where one might want to add extra axioms. These could involve ordering properties, or algebraic properties like those of an Abelian group, and so on. These properties would be realised in any representation as the corresponding type of object in the topos τ_ϕ. Thus Abelian group axioms mean that $\mathcal{T}_\phi$ is an Abelian-group object in τ_ϕ; total-ordering axioms for the time-type $\mathcal{T}$ mean that $\mathcal{T}_\phi$ is a totally-ordered object in τ_ϕ, and so on.

As a rather interesting extension of this idea, one could have a space-time ground type symbol M, but then add the axioms for a partial ordering. In that case, M_ϕ would be a poset-object in τ_ϕ, which could be interpreted physically as the τ_ϕ-analogue of a causal set [57].

6 The Category of Complex Systems

In one sense, there is only one true 'system', and that is the universe as a whole. Concomitantly, there is just one local language, and one topos. However, in practice, the universe is divided conceptually into portions that are sufficiently simple to be amenable to theoretical discussion. Clearly, this division is not unique, but it must be such that the coupling between portions is weak enough that, to a good approximation, their theoretical models can be studied in isolation from each other.[60] Such an essentially isolated portion of the universe is called a 'sub-system'. By an abuse of language, sub-systems of the universe are usually called 'systems' (so that the universe as a whole is one super-system), and then we can talk about 'sub-systems' of these systems; or 'composites' of them; or sub-systems of the composite systems, and so on [41].

To develop these ideas further we need mathematical control over the systems of interest, and their interrelations. To this end, we start by focussing on some collection, **Sys**, of physical systems to which a particular theory-type is deemed to be applicable. For example, we could consider a collection of systems that are to be discussed using the methodology of crowd mechanics; or systems to be discussed using standard quantum theory; or whatever. For completeness, we require that every sub-system of a system in **Sys** is itself a member of **Sys**, as is every composite of members of **Sys**.

We will assume that the systems in **Sys** are all associated with local languages of the type discussed in [38], and that they all have the *same* set of ground symbols which, for the purposes of the present discussion, we take to

---

[60] We remark here that we are actually trying to represent that system by things in it, which the Incompleteness Theorems by Gödel and later authors show is not possible in full. So, it is not a surprise that we should have multiple competing theories and theories that are mutually inconsistent.

be just Σ and $\mathcal{R}$. It follows that the languages $\mathcal{TL}(S)$, $S \in \mathbf{Sys}$, differ from each other only in the set of function symbols $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$; i.e., the set of *physical quantities.*

As a simple example of the system-dependence of the set of function symbols let system S_1 be an agent moving in one dimension, and let the set of physical quantities be $F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) = \{x, p, H\}$. In the language $\mathcal{TL}(S_1)$, these function-symbols represent the position, momentum, and energy of the system respectively. On the other hand, if S_2 is an agent moving in three dimensions, then in the language $\mathcal{TL}(S_2)$ we could have $F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big) = \{x, y, z, p_x, p_y, p_z, H\}$ to allow for three-dimensional position and momentum. Or, we could decide to add angular momentum as well, to give the set $F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big) = \{x, y, z, p_x, p_y, p_z, J_x, J_y, J_z, H\}$.

6.1 The Category Sys

6.1.1 The Arrows and Translations for the Disjoint Sum $S_1 \sqcup S_2$

The use of local languages is central to our overall topos scheme, and therefore we need to understand, in particular, (i) the relation between the languages $\mathcal{TL}(S_1)$ and $\mathcal{TL}(S_2)$ if S_1 is a sub-system of S_2; and (ii) the relation between $\mathcal{TL}(S_1)$, $\mathcal{TL}(S_2)$ and $\mathcal{TL}(S_1 \otimes S_2)$, where $S_1 \otimes S_2$ denotes the composite of systems S_1 and S_2 [41].

These discussions can be made more precise by regarding **Sys** as a category whose objects are crowd systems. The morphisms in **Sys** need to cover two basic types of relation: (i) that between S_1 and S_2 if S_1 is a 'sub-system' of S_2; and (ii) that between a composite crowd system, $S_1 \otimes S_2$, and its constituent systems, S_1 and S_2.

This may seem straightforward but, in fact, care is needed since although the idea of a 'sub-system' seems intuitively clear, it is hard to give a physically acceptable definition that is universal. However, some insight into this idea can be gained by considering its meaning in crowd mechanics. This is very relevant for the general scheme since one of our main goals is to make all theories 'look' like crowd mechanics in the appropriate topos.

To this end, let S_1 and S_2 be classical systems whose state spaces are the symplectic manifolds $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. If S_1 is deemed to be a sub-system of S_2, it is natural to require that $\mathcal{S}_1$ is a *sub-manifold* of $\mathcal{S}_2$, i.e., $\mathcal{S}_1 \subseteq \mathcal{S}_2$. However, this condition cannot be used as a *definition* of a 'sub-system' since the converse may not be true: i.e., if $\mathcal{S}_1 \subseteq \mathcal{S}_2$, this does not necessarily mean that, from a physical perspective, S_1 could, or would, be said to be a sub-system of S_2.

On the other hand, there are situations where being a sub-manifold clearly *does* imply being a physical sub-system. For example, suppose the state space $\mathcal{S}$ of a system S is a disconnected manifold with two components $\mathcal{S}_1$ and $\mathcal{S}_2$, so that $\mathcal{S}$ is the disjoint union, $\mathcal{S}_1 \coprod \mathcal{S}_2$, of the sub-manifolds $\mathcal{S}_1$ and $\mathcal{S}_2$. Then it seems physically appropriate to say that the system S itself is disconnected, and to write $S = S_1 \sqcup S_2$ where the symplectic manifolds that represent the sub-systems S_1 and S_2 are $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.

One reason why it is reasonable to call S_1 and S_2 'sub-systems' in this particular situation is that any continuous dynamical evolution of a state point in $\mathcal{S} \simeq \mathcal{S}_1 \sqcup \mathcal{S}_2$ will always lie in either one component or the other. This suggests that perhaps, in general, a necessary condition for a sub-manifold $\mathcal{S}_1 \subseteq \mathcal{S}_2$ to represent a physical sub-system is that the dynamics of the system S_2 must be such that $\mathcal{S}_1$ is mapped into itself under the dynamical evolution on $\mathcal{S}_2$; in other words, $\mathcal{S}_1$ is a *dynamically-invariant* sub-manifold of $\mathcal{S}_2$. This correlates with the idea mentioned earlier that sub-systems are weakly-coupled with each other.

However, such a dynamical restriction is not something that should be coded into the languages, $\mathcal{TL}(S_1)$ and $\mathcal{TL}(S_2)$: rather, the dynamics is to be associated with the *representation* of these languages in the appropriate topoi.

Still, this caveat does not apply to the disjoint sum $S_1 \sqcup S_2$ of two systems S_1, S_2, and we will assume that, in general, (i.e., not just in crowd mechanics) it is legitimate to think of S_1 and S_2 as being sub-systems of $S_1 \sqcup S_2$; something that we indicate by defining morphisms $i_1 : S_1 \to S_1 \sqcup S_2$, and $i_2 : S_2 \to S_1 \sqcup S_2$ in **Sys** [41].

To proceed further it is important to understand the connection between the putative morphisms in the category **Sys**, and the 'translations' of the associated languages. The first step is to consider what can be said about the relation between $\mathcal{TL}(S_1 \sqcup S_2)$, and $\mathcal{TL}(S_1)$ and $\mathcal{TL}(S_2)$. All three languages share the same ground-type symbols, and so what we are concerned with is the relation between the function symbols of signature $\Sigma \to \mathcal{R}$ in these languages.

By considering what is meant intuitively by the disjoint sum, it seems plausible that each physical quantity for the system $S_1 \sqcup S_2$ produces a physical quantity for S_1, and another one for S_2. Conversely, specifying a pair of physical quantities—one for S_1 and one for S_2—gives a physical quantity for $S_1 \sqcup S_2$. In other words,

$$F_{\mathcal{TL}(S_1 \sqcup S_2)}(\Sigma, \mathcal{R}) \simeq F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) \times F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) \tag{56}$$

However, it is important not to be too dogmatic about statements of this type since in non-classical theories new possibilities can arise that are counter to intuition.

Associated with (56) are the maps $\mathcal{TL}(i_1) : F_{\mathcal{TL}(S_1 \sqcup S_2)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R})$ and $\mathcal{TL}(i_2) : F_{\mathcal{TL}(S_1 \sqcup S_2)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R})$, defined as the projection maps of the product. In the theory of local languages, these transformations are essentially *translations* [28] of $\mathcal{TL}(S_1 \sqcup S_2)$ in $\mathcal{TL}(S_1)$ and $\mathcal{TL}(S_2)$ respectively; a situation that we denote $\mathcal{TL}(i_1) : \mathcal{TL}(S_1 \sqcup S_2) \to \mathcal{TL}(S_1)$, and $\mathcal{TL}(i_2) : \mathcal{TL}(S_1 \sqcup S_2) \to \mathcal{TL}(S_2)$.

To be more precise, these operations are translations if, taking $\mathcal{TL}(i_1)$ as the explanatory example, the map $\mathcal{TL}(i_1) : F_{\mathcal{TL}(S_1 \sqcup S_2)}(\Sigma, \mathcal{R})$

$\to F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big)$ is supplemented with the following map from the ground symbols of $\mathcal{TL}(S_1 \sqcup S_2)$ to those of $\mathcal{TL}(S_1)$ [41]:

$$\mathcal{TL}(i_1)(\Sigma) := \Sigma, \tag{57}$$

$$\mathcal{TL}(i_1)(\mathcal{R}) := \mathcal{R}, \tag{58}$$

$$\mathcal{TL}(i_1)(1) := 1, \tag{59}$$

$$\mathcal{TL}(i_1)(\Omega) := \Omega. \tag{60}$$

Such a translation map is then extended to all type symbols using the definitions

$$\mathcal{TL}(i_1)(T_1 \times T_2 \times \cdots \times T_n) = \mathcal{TL}(i_1)(T_1) \times \mathcal{TL}(i_1)(T_2) \times \cdots \times \mathcal{TL}(i_1)(T_n),$$
$$\mathcal{TL}(i_1)(PT) = P[\mathcal{TL}(i_1)(T)],$$

for all finite n and all type symbols $T, T_1, T_2, \ldots, T_n$. This, in turn, can be extended inductively to all terms in the language. Thus, in our case, the translations act trivially on all the type symbols.

Motivated by this argument we now turn everything around and, in general, *define* a morphism $j : S_1 \to S$ in the category **Sys** to mean that there is some *physically meaningful* way of transforming the physical quantities in S to physical quantities in S_1. If, for any pair of systems S_1, S there is more than one such transformation, then there will be more than one morphism from S_1 to S [41].

To make this more precise, let **Loc** denote the collection of all (small) local languages. This is a category whose objects are the local languages, and whose morphisms are translations between languages. Then our basic assumption is that the association $S \mapsto \mathcal{TL}(S)$ is a covariant functor from **Sys** to $\mathbf{Loc}^{\mathrm{op}}$, which we denote as $\mathcal{L} : \mathbf{Sys} \to \mathbf{Loc}^{\mathrm{op}}$.[61]

6.1.2 The Arrows and Translations for the Composite System $S_1 \otimes S_2$

Let us now consider the composition $S_1 \otimes S_2$ of a pair of systems. In the case of crowd mechanics, if $\mathcal{S}_1$ and $\mathcal{S}_2$ are the symplectic manifolds that represent the systems S_1 and S_2 respectively, then the manifold that represents the composite system is the Cartesian product $\mathcal{S}_1 \times \mathcal{S}_2$. This is distinguished by the existence of the two projection functions $pr_1 : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathcal{S}_1$ and $pr_2 : \mathcal{S}_1 \times \mathcal{S}_2 \to \mathcal{S}_2$ [41].

It seems reasonable to impose the same type of structure on **Sys**: i.e., to require there to be morphisms $p_1 : S_1 \otimes S_2 \to S_1$ and $p_2 : S_1 \otimes S_2 \to S_2$ in **Sys**. However, bearing in mind the definition above, these morphisms p_1, p_2 exist iff there are corresponding translations $\mathcal{TL}(p_1) : \mathcal{TL}(S_1) \to \mathcal{TL}(S_1 \otimes S_2)$, and $\mathcal{TL}(p_2) : \mathcal{TL}(S_2) \to \mathcal{TL}(S_1 \otimes S_2)$. But there *are* such translations: for if

---

[61] The combination of a pair of morphisms in **Sys** exists in so far as the associated translations can be combined.

A_1 is a physical quantity for system S_1, then $\mathcal{TL}(p_1)(A_1)$ can be defined as that same physical quantity, but now regarded as pertaining to the combined system $S_1 \otimes S_2$; and analogously for system S_2. For example, if A is the energy of agent 1, then we can talk about this energy in the combination of a pair of agents. We will denote this translated quantity, $\mathcal{TL}(p_1)(A_1)$, by $A_1 \otimes 1$.[62]

The definitions above of the basic morphisms suggest that we might also want to impose the following conditions [41]:

1. The morphisms $i_1 : S_1 \to S_1 \sqcup S_2$, and $i_2 : S_2 \to S_1 \sqcup S_2$ are *monic* in **Sys**.
2. The morphisms $p_1 : S_1 \otimes S_2 \to S_1$ and $p_2 : S_1 \otimes S_2 \to S_2$ are *epic* morphisms in **Sys**.[63]

6.1.3 The Concept of 'Isomorphic' Systems

We also need to decide what it means to say that two systems S_1 and S_2 are *isomorphic*, to be denoted $S_1 \simeq S_2$. As with the concept of sub-system, the notion of isomorphism is to some extent a matter of definition rather than obvious physical structure, albeit with the expectation that isomorphic systems in **Sys** will correspond to isomorphic local languages, and be represented by isomorphic mathematical objects in any concrete realization of the axioms: for example, by isomorphic symplectic manifolds in crowd mechanics.

To a considerable extent, the physical meaning of 'isomorphism' depends on whether one is dealing with actual physical systems, or idealisations of them. For example, an electron confined in a box in Cambridge is presumably isomorphic to one confined in the same type of box in London, although they are not the same physical system. On the other hand, when a lecturer says "Consider an electron trapped in a box....", he/she is referring to an idealized system.

One could, perhaps, say that an idealized system is an *equivalence class* (under isomorphisms) of real systems, but even working only with idealizations does not entirely remove the need for the concept of isomorphism.

For example, in classical mechanics, consider the (idealized) system S of an agent moving in a box, and let 1 denote the 'trivial system' that consists of just a single point with no internal or external degrees of freedom. Now consider the system $S \otimes 1$. In classical mechanics this is represented by the symplectic manifold $\mathcal{S} \times \{*\}$, where $\{*\}$ is a single point, regarded as a zero-dimensional manifold. However, $\mathcal{S} \times \{*\}$ is isomorphic to the manifold $\mathcal{S}$, and it is clear physically that the system $S \otimes 1$ is isomorphic to the system S. On the other hand, one cannot say that $S \otimes 1$ is literally *equal* to S, so the concept of 'isomorphism' needs to be maintained.

---

[62] We do not postulate any simple relation between $F_{\mathcal{TL}(S_1 \otimes S_2)}(\Sigma, \mathcal{R})$ and $F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R})$ and $F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R})$; i.e., there is no analogue of (56) for combinations of systems.

[63] However, we do not require that $S_1 \cup S_2$ and $S_1 \otimes S_2$ are the co-product and product, respectively, of S_1 and S_2 in the category **Sys**.

One thing that *is* clear is that if $S_1 \simeq S_2$ then $F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) \simeq F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big)$, and if any other non-empty sets of function symbols are present, then they too must be isomorphic.

Note that when introducing a trivial system, 1, it necessary to specify its local language, $\mathcal{TL}(1)$. The set of function symbols $F_{\mathcal{TL}(1)}\big(\Sigma, \mathcal{R}\big)$ is not completely empty since, in crowd mechanics, one does have a preferred physical quantity, which is just the number 1. If one asks what is meant in general by the 'number 1' the answer is not trivial since, in the reals $\mathbb{R}$, the number 1 is the multiplicative identity. It would be possible to add the existence of such a unit to the axioms for $\mathcal{R}$ but this would involve introducing a multiplicative structure and we do not know if there might be physically interesting topos representations that do not have this feature.

For the moment then, we will say that the trivial system has just a single physical quantity, which in crowd mechanics translates to the number 1. More generally, for the language $\mathcal{TL}(1)$ we specify that $F_{\mathcal{TL}(1)}\big(\Sigma, \mathcal{R}\big) := \{I\}$, i.e., $F_{\mathcal{TL}(1)}\big(\Sigma, \mathcal{R}\big)$ has just a single element, I, say. Furthermore, we add the axiom [41]

$$: \forall \tilde{s}_1 \forall \tilde{s}_2, I(\tilde{s}_1) = I(\tilde{s}_2),$$

where $\tilde{s}_1$ and $\tilde{s}_2$ are variables of type Σ. In fact, it seems natural to add such a trivial quantity to the language $\mathcal{TL}(S)$ for *any* system S, and from now on we will assume that this has been done.

A related issue is that, in crowd mechanics, if A is a physical quantity, then so is rA for any $r \in \mathbb{R}$. This is because the set of classical quantities $A_s : \Sigma_s \to \mathcal{R}_s \simeq \mathbb{R}$ forms a *ring* whose structure derives from the ring structure of $\mathbb{R}$. It would be possible to add ring axioms for $\mathcal{R}$ to the language $\mathcal{TL}(S)$, but we think this is too strong because it fails in quantum theory [40].

If desired, an '*empty*' system, 0, can be added too, with $F_{\mathcal{TL}(0)}\big(\Sigma, \mathcal{R}\big) := \emptyset$. This, so called, 'pure language', $\mathcal{TL}(0)$, is an initial object in the category **Loc**.

6.1.4 An Axiomatic Formulation of the Category Sys

Here we summarize the list of axioms for a category **Sys** [41]:

1. The collection **Sys** is a small category where (i) the objects are the systems of interest (or, if desired, isomorphism classes of such systems); and (ii) the morphisms are defined as above.

   Thus the fundamental property of a morphism $j : S_1 \to S$ in **Sys** is that it induces, and is essentially *defined by*, a translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$. Physically, it corresponds to the physical quantities for system S being 'pulled-back' to give physical quantities for system S_1.

   Arrows of particular interest are those associated with 'sub-systems' and 'composite systems', as discussed above.

2. The axioms for a category are satisfied because:
   (a) Physically, the ability to form composites of morphisms follows from the concept of 'pulling-back' physical quantities. From a mathematical perspective, if $j : S_1 \to S_2$ and $k : S_2 \to S_3$, then the translations give functions $\mathcal{TL}(j) : F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big) \to F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big)$ and $\mathcal{TL}(k) : F_{\mathcal{TL}(S_3)}\big(\Sigma, \mathcal{R}\big) \to F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big)$. Then clearly $\mathcal{TL}(j) \circ \mathcal{TL}(k) : F_{\mathcal{TL}(S_3)}\big(\Sigma, \mathcal{R}\big) \to F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big)$, and this can thought of as the translation corresponding to the morphism $k \circ j : S_1 \to S_3$.
   The associativity of the law of morphism combination can be proved in a similar way.
   (b) We add by hand a special morphism $\mathrm{id}_S : S \to S$ which is defined to correspond to the translation $\mathcal{TL}(id_S)$ that is given by the identity map on $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$. Clearly, $\mathrm{id}_S : S \to S$ acts an an identity morphism should.
3. For any pair of systems S_1, S_2, there is a *disjoint sum*, denoted $S_1 \sqcup S_2$. The disjoint sum has the following properties:
   (a) For all systems S_1, S_2, S_3 in **Sys**:
   $$(S_1 \sqcup S_2) \sqcup S_3 \simeq S_1 \sqcup (S_2 \sqcup S_3).$$
   (b) For all systems S_1, S_2 in **Sys**:
   $$S_1 \sqcup S_2 \simeq S_2 \sqcup S_1.$$
   (c) There are morphisms in **Sys**:
   $$i_1 : S_1 \to S_1 \sqcup S_2 \ \text{ and } i_2 : S_2 \to S_1 \sqcup S_2,$$
   that are associated with translations, which are associated with the decomposition
   $$F_{\mathcal{TL}(S_1 \sqcup S_2)}\big(\Sigma, \mathcal{R}\big) \simeq F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) \times F_{\mathcal{TL}(S_2)}\big(\Sigma, \mathcal{R}\big). \tag{61}$$
   We assume that if S_1, S_2 belong to **Sys**, then **Sys** also contains $S_1 \sqcup S_2$.
4. For any given pair of systems S_1, S_2, there is a *composite* system in **Sys**, denoted $S_1 \otimes S_2$, with the following properties:
   (a) For all systems S_1, S_2, S_3 in **Sys**:
   $$(S_1 \otimes S_2) \otimes S_3 \simeq S_1 \otimes (S_2 \otimes S_3). \tag{62}$$
   (b) For all systems S_1, S_2 in **Sys**:
   $$S_1 \otimes S_2 \simeq S_2 \otimes S_1. \tag{63}$$
   (c) There are morphisms in **Sys**:
   $$p_1 : S_1 \otimes S_2 \to S_1 \text{ and } p_2 : S_1 \otimes S_2 \to S_2,$$
   that are associated with translations.

We assume that if S_1, S_2 belong to **Sys**, then **Sys** also contains the composite system $S_1 \otimes S_2$.

5. It seems physically reasonable to add the axiom

$$(S_1 \sqcup S_2) \otimes S \simeq (S_1 \otimes S) \sqcup (S_2 \otimes S) \tag{64}$$

for all systems S_1, S_2, S. However, physical intuition can be a dangerous thing, and so, as with most of these axioms, we are not dogmatic, and feel free to change them as new insights emerge.

6. There is a trivial system, 1, such that for all systems S, we have

$$S \otimes 1 \simeq S \simeq 1 \otimes S \tag{65}$$

7. It may be convenient to postulate an 'empty system', 0, with the properties

$$S \otimes 0 \simeq 0 \otimes S \simeq 0, \qquad S \sqcup 0 \simeq 0 \sqcup S \simeq S \qquad \text{for all systems } S.$$

Within the meaning given to morphisms in **Sys**, 0 is a *terminal object* in **Sys**. This is because the empty set of function symbols of signature $\Sigma \to \mathcal{R}$ is a subset of any other set of function symbols of this signature.

It might seem tempting to postulate that composition laws are well-behaved with respect to morphisms. Namely, if $j : S_1 \to S_2$, then, for any S, there is a morphism $S_1 \otimes S \to S_2 \otimes S$ and a morphism $S_1 \sqcup S \to S_2 \sqcup S$.[64]

In the case of the disjoint sum, such a morphism can be easily constructed using (61). First split the function symbols in $F_{\mathcal{TL}(S_1 \sqcup S)}(\Sigma, \mathcal{R})$ into $F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) \times F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$ and the function symbols in $F_{\mathcal{TL}(S_2 \sqcup S)}(\Sigma, \mathcal{R})$ into $F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) \times F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$. Since there is a morphism $j : S_1 \to S_2$, there is a translation $\mathcal{TL}(j) : \mathcal{TL}(S_2) \to \mathcal{TL}(S_1)$, given by a mapping $\mathcal{TL}(j) : F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R})$. Clearly, then there is also a mapping $\mathcal{TL}(j) \times \mathcal{TL}(id_S) : F_{\mathcal{TL}(S_2)}(\Sigma, \mathcal{R}) \times F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) \times F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$, i.e., a translation between $\mathcal{TL}(S_2 \sqcup S)$ and $\mathcal{TL}(S_1 \sqcup S)$. Since we assume that there is a morphism in **Sys** whenever there is a translation (in the opposite direction), there is indeed a morphism $S_1 \sqcup S \to S_2 \sqcup S$.

In the case of the composition, however, this would require a translation $\mathcal{TL}(S_2 \otimes S) \to \mathcal{TL}(S_1 \otimes S)$, and this cannot be done in general since we have no *prima facie* information about the set of function symbols $F_{\mathcal{TL}(S_2 \otimes S)}(\Sigma, \mathcal{R})$. However, if we restrict the morphisms in **Sys** to be those associated with subsystems, combination of systems, and compositions of such morphisms, then it is easy to see that the required translations exist (the proof of this makes essential use of (64)).

---

[64] A more accurate way of capturing this idea is to say that the operation **Sys** × **Sys** → **Sys** in which

$$\langle S_1, S_2 \rangle \mapsto S_1 \otimes S_2 \tag{66}$$

is a *bi-functor* from **Sys** × **Sys** to **Sys**.

If we make this restriction of morphisms, then the axioms (63), (65–66), mean that, essentially, **Sys** has the structure of a *symmetric monoidal* category in which the monoidal product operation is '$\otimes$', and the left and right unit object is 1. There is also a monoidal structure associated with the disjoint sum '$\sqcup$', with 0 as the unit object [41].

We say 'essentially' because in order to comply with all the axioms of a monoidal category, **Sys** must satisfy certain additional *coherence axioms.* However, from a physical perspective these are very plausible statements about (i) how the unit object 1 intertwines with the $\otimes$-operation; (ii) how the null object intertwines with the $\sqcup$-operation; and (iii) certain properties of quadruple products (and disjoint sums) of systems.

It might be helpful at this point to give a simple example of a category **Sys**. Let S denote an agent that moves in three dimensions, and let us suppose that S has no sub-systems other than the trivial system 1. Then $S \otimes S$ is defined to be a pair of agents moving in three dimensions, and so on. Thus the objects in our category are 1, S, $S \otimes S$, ..., $S \otimes S \otimes \cdots S$ ... where the '$\otimes$' operation is formed any finite number of times.

At this stage, the only morphisms are those that are associated with the constituents of a composite crowd system. However, we could contemplate adding to the systems the disjoint sum $S \sqcup (S \otimes S)$ which is a system that is either one agent or two agents (but, of course, not both at the same time). And, clearly, we could extend this to $S \sqcup (S \otimes S) \sqcup (S \otimes S \otimes S)$, and so on. Each of these disjoint sums comes with its own morphisms, as explained above.

6.2 Representations of Sys in Topoi

We assume that all the systems in **Sys** are to be treated with the same theory type. We also assume that systems in **Sys** with the *same* language are to be represented in the same topos. Then we can define:[65] [41] A *topos realization* of **Sys** is an association, ϕ, to each system S in **Sys**, of a triple $\phi(S) = \langle \rho_{\phi,S}, \mathcal{TL}(S), \tau_\phi(S) \rangle$ where:

(i) $\tau_\phi(S)$ is the topos in which the theory-type applied to system S is to be realised.
(ii) $\mathcal{TL}(S)$ is the local language in **Loc** that is associated with S. This is not dependent on the realization ϕ.
(iii) $\rho_{\phi,S}$ is a representation of the local language $\mathcal{TL}(S)$ in the topos $\tau_\phi(S)$. As a more descriptive piece of notation we write $\rho_{\phi,S} : \mathcal{TL}(S) \rightsquigarrow \tau_\phi(S)$. The key part of this representation is the map

$$\rho_{\phi,S} : F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \to \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S},$$

---

[65] As emphasised already, the association $S \mapsto \mathcal{TL}(S)$ is generally not one-to-one: i.e., many systems may share the same language. Thus, when we come discuss the representation of the language $\mathcal{TL}(S)$ in a topos, the extra information about the system S is used in fixing the representation.

> where $\Sigma_{\phi,S}$ and $\mathcal{R}_{\phi,S}$ are the state object and quantity-value object, respectively, of the representation ϕ in the topos $\tau_\phi(S)$. As a convenient piece of notation we write $A_{\phi,S} := \rho_{\phi,S}(A)$ for all $A \in F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$.

Now, if $j : S_1 \to S$ is a morphism in **Sys**, then there is a translation morphism $\mathcal{L}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$. Thus we have the beginnings of a commutative diagram [41]

$$\begin{tikzcd}
S_1 \arrow[r, "\phi"] \arrow[d, "j"'] & \langle \rho_{\phi,S_1}, \mathcal{TL}(S_1), \tau_\phi(S_1)\rangle \\
S \arrow[r, "\phi"'] & \langle \rho_{\phi,S}, \mathcal{TL}(S), \tau_\phi(S)\rangle \arrow[u, "? \times \mathcal{L}(j) \times ?"']
\end{tikzcd} \tag{67}$$

However, to be useful, the morphism on the right hand side of this diagram should refer to some relation between (i) the topoi $\tau_\phi(S_1)$ and $\tau_\phi(S)$; and (ii) the realizations $\rho_{\phi,S_1} : \mathcal{TL}(S_1) \rightsquigarrow \tau_\phi(S_1)$ and $\rho_{\phi,S} : \mathcal{TL}(S) \rightsquigarrow \tau_\phi(S)$: this is the significance of the two '?' symbols in the morphism written '$? \times \mathcal{TL}(j) \times ?$'.

Indeed the above definition says nothing about relations between the topoi representations of different systems in **Sys**. We are particularly interested in the situation where there are two different systems S_1 and S with a morphism $j : S_1 \to S$ in **Sys**.

We know that the morphism j is associated with a translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$, and an attractive possibility, therefore, would be to seek, or postulate, a 'covering' map $\phi(\mathcal{TL}(j)) : \texttt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \texttt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$ to be construed as a topos representation of the translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$, and hence of the morphism $j : S_1 \to S$ in **Sys**.

This raises the questions of what properties these 'translation representations' should possess in order to justify saying that they 'cover' the translations. A minimal requirement is that if $k : S_2 \to S_1$ and $j : S_1 \to S$, then the map $\phi(\mathcal{TL}(j \circ k)) : \texttt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \texttt{Mor}\tau_\phi(S_2)\Sigma_{\phi,S_2}\mathcal{R}_{\phi,S_2}$ factorises as

$$\phi(\mathcal{TL}(j \circ k)) = \phi(\mathcal{TL}(k)) \circ \phi(\mathcal{TL}(j)). \tag{68}$$

We also require that

$$\phi(\mathcal{TL}(id_S)) = id : \texttt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \texttt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \tag{69}$$

for all systems S.

The conditions (68) and (69) seem eminently plausible. They are not particularly strong. A far more restrictive axiom would be to require the following diagram to commute [41]:

$$\begin{tikzcd}
F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \arrow[r, "\rho_{\phi,S}"] \arrow[d, "\mathcal{TL}(j)"'] & \mathtt{Mor}_{\tau_\phi(S)} \Sigma_{\phi,S} \mathcal{R}_{\phi,S} \arrow[d, "\phi(\mathcal{TL}(j))"] \\
F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) \arrow[r, "\rho_{\phi,S_1}"'] & \mathtt{Mor}_{\tau_\phi(S_1)} \Sigma_{\phi,S_1} \mathcal{R}_{\phi,S_1}
\end{tikzcd} \tag{70}$$

At first sight, this requirement seems very appealing. However, caution is needed when postulating 'axioms' for a theoretical structure in physics. It is easy to get captivated by the underlying mathematics and to assume, erroneously, that what is mathematically elegant is necessarily useful in the physical theory.

The translation $\phi(\mathcal{TL}(j))$ maps a morphism from $\Sigma_{\phi,S}$ to $\mathcal{R}_{\phi,S}$ to a morphism from Σ_{ϕ,S_1} to $\mathcal{R}_{\phi,S_1}$. Intuitively, if Σ_{ϕ,S_1} is a 'much larger' object than $\Sigma_{\phi,S}$ (they lie in different topoi, so no direct comparison is available), the translation can only be 'faithful' on some part of Σ_{ϕ,S_1} that can be identified with (the 'image' of) $\Sigma_{\phi,S}$. A concrete example of this will show up in the treatment of composite quantum systems. As one might expect, a form of entanglement plays a role here.

6.3 Crowd Mechanics in Sys

Constructing maps $\phi(\mathcal{TL}(j)) : \mathtt{Mor}_{\tau_\phi(S)} \Sigma_{\phi,S} \mathcal{R}_{\phi,S} \to \mathtt{Mor}_{\tau_\phi(S_1)} \Sigma_{\phi,S_1} \mathcal{R}_{\phi,S_1}$ is likely to be complicated when $\tau_\phi(S)$ and $\tau_\phi(S_1)$ are different topoi, and so we begin with the example of crowd mechanics, where the topos is always **Set**.

In general, we are interested in the relation(s) between the representations $\rho_{\phi,S_1} : \mathcal{TL}(S_1) \rightsquigarrow \tau_\phi(S_1)$ and $\rho_{\phi,S} : \mathcal{TL}(S) \rightsquigarrow \tau_\phi(S)$ that is associated with a morphism $j : S_1 \to S$ in **Sys**. In crowd mechanics, we only have to study the relation between the representations $\rho_{\sigma,S_1} : \mathcal{TL}(S_1) \rightsquigarrow$ **Set** and $\rho_{\sigma,S} : \mathcal{TL}(S) \rightsquigarrow$ **Set**.

Let us summarise what we have said so far (with σ denoting the **Set**-realization of crowd mechanics):

1. For any system S in **Sys**, a representation $\rho_{\sigma,S} : \mathcal{TL}(S) \rightsquigarrow$ **Set** consists of the following ingredients [41]:
   (a) The ground symbol Σ is represented by a symplectic manifold, $\Sigma_{\sigma,S} := \rho_{\sigma,S}(\Sigma)$, that serves as the classical state space.
   (b) For all systems S, the ground symbol $\mathcal{R}$ is represented by the real numbers $\mathbb{R}$, i.e., $\mathcal{R}_{\sigma,S} = \mathbb{R}$, where $\mathcal{R}_{\sigma,S} := \rho_{\sigma,S}(\mathcal{R})$.

   (c) Each function symbol $A : \Sigma \to \mathcal{R}$ in $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$ is represented by a function $A_{\sigma,S} = \rho_{\sigma,S}(A) : \Sigma_{\sigma,S} \to \mathbb{R}$ in the set of functions[66] $C(\Sigma_{\sigma,S}; \mathbb{R})$.

2. The trivial system is mapped to a singleton set $\{*\}$ (viewed as a zero-dimensional symplectic manifold):

$$\Sigma_{\sigma,1} := \{*\}.$$

   The empty system is represented by the empty set:

$$\Sigma_{\sigma,0} := \emptyset.$$

3. Propositions about the system S are represented by (Borel) subsets of the state space $\Sigma_{\sigma,S}$.
4. The composite system $S_1 \otimes S_2$ is represented by the Cartesian product $\Sigma_{\sigma,S_1} \times \Sigma_{\sigma,S_2}$; i.e.,

$$\Sigma_{\sigma,\, S_1 \otimes S_2} \simeq \Sigma_{\sigma,S_1} \times \Sigma_{\sigma,S_2}. \tag{71}$$

   The disjoint sum $S_1 \sqcup S_2$ is represented by the disjoint union $\Sigma_{\sigma,S_1} \coprod \Sigma_{\sigma,S_2}$; i.e.,

$$\Sigma_{\sigma,S_1 \sqcup S_2} \simeq \Sigma_{\sigma,S_1} \coprod \Sigma_{\sigma,S_2}.$$

5. Let $j : S_1 \to S$ be a morphism in **Sys**. Then
   (a) There is a translation map $\mathcal{TL}(j) : F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big) \to F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big)$.
   (b) There is a symplectic function $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$ from the symplectic manifold Σ_{σ,S_1} to the symplectic manifold $\Sigma_{\sigma,S}$.

The existence of this function $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$ follows directly from the properties of sub-systems and composite systems in crowd mechanics. These properties of the morphisms stem from the fact that the linguistic function symbols in $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$ are represented by real-valued functions in $C(\Sigma_{\sigma,S}, \mathbb{R})$. Thus we can write $\rho_{\sigma,S} : F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big) \to C(\Sigma_{\sigma,S}, \mathbb{R})$, and similarly $\rho_{\sigma,S_1} : F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) \to C(\Sigma_{\sigma,S_1}, \mathbb{R})$. The diagram in (70) now becomes [41]

$$\begin{tikzcd}
F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big) \arrow[r, "\rho_{\sigma,S}"] \arrow[d, "\mathcal{TL}(j)"'] & C(\Sigma_{\sigma,S}, \mathbb{R}) \arrow[d, "\sigma(\mathcal{TL}(j))"] \\
F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) \arrow[r, "\rho_{\sigma,S_1}"'] & C(\Sigma_{\sigma,S_1}, \mathbb{R})
\end{tikzcd} \tag{72}$$

and, therefore, the question of interest is if there is a 'translation representation' function $\sigma(\mathcal{TL}(j)) : C(\Sigma_{\sigma,S}, \mathbb{R}) \to C(\Sigma_{\sigma,S_1}, \mathbb{R})$ so that this diagram commutes.

---

[66] In practice, these functions are required to be measurable with respect to the Borel structures on the symplectic manifold Σ_s and $\mathbb{R}$.

Now, as stated above, a physical quantity, A, for the system S is represented in crowd mechanics by a real-valued function $A_{\sigma,S} = \rho_{\sigma,S}(A) : \Sigma_{\sigma,S} \to \mathbb{R}$. Similarly, the representation of $\mathcal{TL}(j)(A)$ for S_1 is given by a function $A_{\sigma,S_1} := \rho_{\sigma,S_1}(A) : \Sigma_{\sigma,S_1} \to \mathbb{R}$. However, in this classical case we also have the function $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$, and it is clear that we can use it to define $[\rho_{\sigma,S_1}(\mathcal{TL}(j)(A)](s) := \rho_{\sigma,S}(A)\big(\sigma(j)(s)\big)$ for all $s \in \Sigma_{\sigma,S_1}$. In other words, we have

$$\big((\mathcal{TL}(j)(A)\big)_{\sigma,S_1} = A_{\sigma,S} \circ \sigma(j).$$

But then it is clear that a translation-representation function $\sigma(\mathcal{TL}(j)) : C(\Sigma_{\sigma,S}, \mathbb{R}) \to C(\Sigma_{\sigma,S_1}, \mathbb{R})$ with the desired property of making (72) commute can be defined by

$$\sigma(\mathcal{TL}(j))(f) := f \circ \sigma(j) \tag{73}$$

for all $f \in C(\Sigma_{\sigma,S}, \mathbb{R})$; i.e., the function $\sigma(\mathcal{TL}(j))(f) : \Sigma_{\sigma,S_1} \to \mathbb{R}$ is the usual pull-back of the function $f : \Sigma_{\sigma,S} \to \mathbb{R}$ by the function $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$.

Thus, in the case of crowd mechanics, the commutative diagram in (67) can be completed to give [41]

$$\begin{tikzcd}
S_1 \arrow[r, "\sigma"] \arrow[d, "j"'] & \langle \rho_{\sigma,S_1}, \mathcal{TL}(S_1), \mathbf{Set} \rangle \\
S \arrow[r, "\sigma"'] & \langle \rho_{\sigma,S}, \mathcal{TL}(S), \mathbf{Set} \rangle \arrow[u, "\sigma(\mathcal{TL}(j)) \times \mathcal{TL}(j) \times id"']
\end{tikzcd} \tag{74}$$

6.3.1 Details of the Translation Representation

We first consider morphisms of the form

$$S_1 \overset{i_1}{\to} S_1 \sqcup S_2 \overset{i_2}{\leftarrow} S_2,$$

from the components S_1, S_2 to the disjoint sum $S_1 \sqcup S_2$. The systems S_1, S_2 and $S_1 \sqcup S_2$ have symplectic manifolds Σ_{σ,S_1}, Σ_{σ,S_2} and $\Sigma_{\sigma,S_1 \sqcup S_2} = \Sigma_{\sigma,S_1} \coprod \Sigma_{\sigma,S_2}$. We write $i := i_1$.

Let S be a classical system. We assume that the function symbols $A \in F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$ in the language $\mathcal{TL}(S)$ are in bijective correspondence with an appropriate subset of the functions $A_{\sigma,S} \in C(\Sigma_{\sigma,S}, \mathbb{R})$.[67]

There is an obvious translation representation. For if $A \in F_{\mathcal{TL}(S_1 \sqcup S_2)}\big(\Sigma, \mathcal{R}\big)$, then since $\Sigma_{\sigma,S_1 \sqcup S_2} = \Sigma_{\sigma,S_1} \coprod \Sigma_{\sigma,S_1}$, the associated function $A_{\sigma,S_1 \sqcup S_2} : \Sigma_{\sigma,S_1 \sqcup S_2} \to \mathbb{R}$ is given by a pair of functions $A_1 \in C(\Sigma_{\sigma,S_1}, \mathbb{R})$ and $A_2 \in C(\Sigma_{\sigma,S_2}, \mathbb{R})$; we write $A_{\sigma,S_1 \sqcup S_2} = \langle A_1, A_2 \rangle$. It is natural to demand that the translation representation $\sigma(\mathcal{TL}(i))(A_{\sigma,S_1 \sqcup S_2})$ is A_1. Note that what is

---

[67] Depending on the setting, one can assume that $F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big)$ contains function symbols corresponding bijectively to measurable, continuous or smooth functions.

essentially being discussed here is the classical-physics representation of the relation (56).

The canonical choice for $\sigma(i)$ is [41]

$$\begin{aligned}\sigma(i) : \Sigma_{\sigma,S_1} &\to \Sigma_{\sigma,S_1\sqcup S_2} = \Sigma_{\sigma,S_1}\coprod\Sigma_{\sigma,S_2}\\ s_1 &\mapsto s_1.\end{aligned}$$

Then the pull-back along $\sigma(i)$,

$$\begin{aligned}\sigma(i)^* : C(\Sigma_{\sigma,S_1\sqcup S_2},\mathbb{R}) &\to C(\Sigma_{\sigma,S_1},\mathbb{R})\\ A_{\sigma,S_1\sqcup S_2} &\mapsto A_{\sigma,S_1\sqcup S_2}\circ\sigma(i),\end{aligned}$$

maps (or 'translates') the topos representative $A_{\sigma,S_1\sqcup S_2} = \langle A_1, A_2\rangle$ of the function symbol $A \in F_{\mathcal{TL}(S_1\sqcup S_2)}\big(\Sigma,\mathcal{R}\big)$ to a real-valued function $A_{\sigma,S_1\sqcup S_2}\circ\sigma(i)$ on Σ_{σ,S_1}. This function is clearly equal to A_1.

We now consider morphisms in **Sys** of the form

$$S_1 \overset{p_1}{\leftarrow} S_1\otimes S_2 \overset{p_2}{\rightarrow} S_2,$$

from the composite classical system $S_1\otimes S_2$ to the constituent systems S_1 and S_2. Here, p_1 signals that S_1 is a constituent of the composite system $S_1\otimes S_2$, likewise p_2. The systems S_1, S_2 and $S_1\otimes S_2$ have symplectic manifolds Σ_{σ,S_1}, Σ_{σ,S_2} and $\Sigma_{\sigma,S_1\otimes S_2} = \Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2}$, respectively; i.e., the state space of the composite system $S_1\otimes S_2$ is the Cartesian product of the state spaces of the components. For typographical simplicity in what follows we denote $p := p_1$.

There is a canonical translation $\mathcal{TL}(p)$ between the languages $\mathcal{TL}(S_1)$ and $\mathcal{TL}(S_1\otimes S_2)$ whose representation is the following. Namely, if A is in $F_{\mathcal{TL}(S_1)}\big(\Sigma,\mathcal{R}\big)$, then the corresponding function $A_{\sigma,S_1}\in C(\Sigma_{\sigma,S_1},\mathbb{R})$ is translated to a function $\sigma(\mathcal{TL}(p))(A_{\sigma,S_1})\in C(\Sigma_{\sigma,S_1\otimes S_2},\mathbb{R})$ such that

$$\sigma(\mathcal{TL}(p))(A_{\sigma,S_1})(s_1,s_2) = A_{\sigma,S_1}(s_1) \qquad \text{for all } (s_1,s_2)\in\Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2}.$$

This natural translation representation is based on the fact that, for the symplectic manifold $\Sigma_{\sigma,S_1\otimes S_2} = \Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2}$, each point $s\in\Sigma_{\sigma,S_1\otimes S_2}$ can be identified with a pair, (s_1,s_2), of points $s_1\in\Sigma_{\sigma,S_1}$ and $s_2\in\Sigma_{\sigma,S_2}$. This is possible since the Cartesian product $\Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2}$ is a product in the categorial sense and hence has projections $\Sigma_{\sigma,S_1}\leftarrow\Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2}\rightarrow\Sigma_{\sigma,S_2}$. Then the translation representation of functions is constructed in a straightforward manner. Thus, let [41]

$$\begin{aligned}\sigma(p) : \Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2} &\to \Sigma_{\sigma,S_1}\\ (s_1,s_2) &\mapsto s_1\end{aligned}$$

be the canonical projection. Then, if $A_{\sigma,S_1}\in C(\Sigma_{\sigma,S_1},\mathbb{R})$, the function

$$A_{\sigma,S_1}\circ\sigma(p)\in C(\Sigma_{\sigma,S_1}\times\Sigma_{\sigma,S_2},\mathbb{R})$$

is such that, for all $(s_1, s_2) \in \Sigma_{\sigma,S_1} \times \Sigma_{\sigma,S_2}$,

$$A_{\sigma,S_1} \circ \sigma(p)(s_1, s_2) = A_{\sigma,S_1}(s_1).$$

Thus we can define

$$\sigma(\mathcal{TL}(p))(A_{\sigma,S_1}) := A_{\sigma,S_1} \circ \sigma(p).$$

Clearly, $\sigma(\mathcal{TL}(p))(A_{\sigma,S_1})$ can be seen as the representation of the function symbol $A \otimes 1 \in F_{\mathcal{TL}(S_1 \otimes S_2)}\big(\Sigma, \mathcal{R}\big)$.

7 General Crowd Dynamics in a General Topos

7.1 The Pull-Back Operations

7.1.1 The Pull-Back of Physical Quantities

Motivated by the above, let us try now to see what can be said about the scheme in general. Basically, what is involved is the topos representation of translations of languages. To be more precise, let $j : S_1 \to S$ be a morphism in **Sys**, so that there is a translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$ defined by the translation function $\mathcal{TL}(j) : F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big) \to F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big)$. Now suppose that the systems S and S_1 are represented in the topoi $\tau_\phi(S)$ and $\tau_\phi(S_1)$ respectively. Then, in these representations, the function symbols of signature $\Sigma \to \mathcal{R}$ in $\mathcal{TL}(S)$ and $\mathcal{TL}(S_1)$ are represented by elements of $\mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S}$ and $\mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$ respectively [41].

Our task is to find a function

$$\phi(\mathcal{TL}(j)) : \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1},$$

that can be construed as the topos representation of the translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$, and hence of the morphism $j : S_1 \to S$ in **Sys**. We are particularly interested in seeing if $\phi(\mathcal{TL}(j))$ can be chosen so that the following diagram commutes (compare with (70)):

$$\begin{tikzcd}
F_{\mathcal{TL}(S)}\big(\Sigma, \mathcal{R}\big) \arrow[r, "\rho_{\phi,S}"] \arrow[d, "\mathcal{TL}(j)"'] & \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \arrow[d, "\phi(\mathcal{TL}(j))"] \\
F_{\mathcal{TL}(S_1)}\big(\Sigma, \mathcal{R}\big) \arrow[r, "\rho_{\phi,S_1}"'] & \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}
\end{tikzcd} \tag{75}$$

However, as has been emphasised already, it is not clear that one should expect to find a function $\phi(\mathcal{TL}(j)) : \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$ with this property. The existence and/or properties of such a function will be dependent on the theory-type, and it seems unlikely that much can be said in general about the diagram (75). Nevertheless, let us see how far we can get in discussing the existence of such a function in general.

Thus, if $\mu \in \texttt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S}$, the critical question is if there is some 'natural' way whereby this morphism can be 'pulled-back' to give an element $\phi(\mathcal{TL}(j))(\mu) \in \texttt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$.

The first pertinent remark is that μ is a morphism in the topos $\tau_\phi(S)$, whereas the sought-for pull-back will be a morphism in the topos $\tau_\phi(S_1)$, and so we need a mechanism for getting from one topos to the other (this problem, of course, does not arise in crowd mechanics since the topos of every representation is always **Set**).

The obvious way of implementing this change of topos is via some *functor*, $\tau_\phi(j)$ from $\tau_\phi(S)$ to $\tau_\phi(S_1)$. Indeed, given such a functor, a morphism $\mu : \Sigma_{\phi,S} \to \mathcal{R}_{\phi,S}$ in $\tau_\phi(S)$ is transformed to the morphism

$$\tau_\phi(j)(\mu) : \tau_\phi(j)(\Sigma_{\phi,S}) \to \tau_\phi(j)(\mathcal{R}_{\phi,S}) \tag{76}$$

in $\tau_\phi(S_1)$.

To convert this to a morphism from Σ_{ϕ,S_1} to $\mathcal{R}_{\phi,S_1}$, we need to supplement (76) with a pair of morphisms $\phi(j), \beta_\phi(j)$ in $\tau_\phi(S_1)$ to get the diagram [41]:

$$\begin{tikzcd}
\tau_\phi(j)(\Sigma_{\phi,S}) \arrow[r, "\tau_\phi(j)(\mu)"] & \tau_\phi(j)(\mathcal{R}_{\phi,S}) \arrow[d, "\beta_\phi(j)"] \\
\Sigma_{\phi,S_1} \arrow[u, "\phi(j)"] & \mathcal{R}_{\phi,S_1}
\end{tikzcd} \tag{77}$$

The pull-back, $\phi(\mathcal{TL}(j))(\mu) \in \texttt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$, with respect to these choices can then be defined as

$$\phi(\mathcal{TL}(j))(\mu) := \beta_\phi(j) \circ \tau_\phi(j)(\mu) \circ \phi(j). \tag{78}$$

It follows that a key part of the construction of a topos representation, ϕ, of **Sys** will be to specify the functor $\tau_\phi(j)$ from $\tau_\phi(S)$ to $\tau_\phi(S_1)$, and the morphisms $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)(\Sigma_{\phi,S})$ and $\beta_\phi(j) : \tau_\phi(j)(\mathcal{R}_{\phi,S}) \to \mathcal{R}_{\phi,S_1}$ in the topos $\tau_\phi(S_1)$. These need to be defined in such a way as to be consistent with a chain of morphisms $S_2 \to S_1 \to S$.

When applied to the representative $A_{\phi,S} : \Sigma_{\phi,S} \to \mathcal{R}_{\phi,S}$ of a physical quantity $A \in F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R})$, the diagram (77) becomes (augmented with the upper half) [41]

$$\begin{tikzcd}
\Sigma_{\phi,S} \arrow[r, "A_{\phi,S}"] \arrow[d, "\tau_\phi(j)"'] & \mathcal{R}_{\phi,S} \arrow[d, "\tau_\phi(j)"] \\
\tau_\phi(j)(\Sigma_{\phi,S}) \arrow[r, "\tau_\phi(j)(A_{\phi,S})"'] & \tau_\phi(j)(\mathcal{R}_{\phi,S}) \arrow[d, "\beta_\phi(j)"] \\
\Sigma_{\phi,S_1} \arrow[u, "\phi(j)"] \arrow[r, "\phi(\mathcal{TL}(j))(A_{\phi,S})"'] & \mathcal{R}_{\phi,S_1}
\end{tikzcd} \tag{79}$$

The commutativity of (75) would then require

$$\phi(\mathcal{TL}(j))(A_{\phi,S}) = (\mathcal{TL}(j)A)_{\phi,S_1} \tag{80}$$

or, in a more expanded notation,

$$\phi(\mathcal{TL}(j)) \circ \rho_{\phi,S} = \rho_{\phi,S_1} \circ \mathcal{TL}(j), \tag{81}$$

where both the left hand side and the right hand side of (81) are mappings from $F_{\mathcal{TL}(S)}\big(\Sigma,\mathcal{R}\big)$ to $\mathtt{Mor}_{\tau_\phi}(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$.

Note that the analogous diagram in crowd mechanics is simply

$$\begin{tikzcd}
\Sigma_{\sigma,S} \arrow[r, "A_{\sigma,S}"] & \mathbb{R} \arrow[d, "id"] \\
\Sigma_{\sigma,S_1} \arrow[u, "\sigma(j)"] \arrow[r, "\sigma(\mathcal{TL}(j))(A_{\sigma,S})"'] & \mathbb{R}
\end{tikzcd} \tag{82}$$

and the commutativity/pull-back condition (80) becomes

$$\sigma(\mathcal{TL}(j))(A_{\sigma,S}) = (\mathcal{TL}(j)A)_{\phi,S_1},$$

which is satisfied by virtue of (73).

It is clear from the above that the morphism $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)(\Sigma_{\phi,S})$ can be viewed as the topos analogue of the map $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$ that arises in crowd mechanics whenever there is a morphism $j : S_1 \to S$.

7.1.2 The Pull-Back of Propositions

More insight can be gained into the nature of the triple $\langle \tau_\phi(j), \phi(j), \beta_\phi(j)\rangle$ by considering the analogous operation for propositions. First, consider a morphism $j : S_1 \to S$ in **Sys** in crowd mechanics. Associated with this there is (i) a translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$; (ii) an associated translation mapping $\mathcal{TL}(j) : F_{\mathcal{TL}(S)}\big(\Sigma,\mathcal{R}\big) \to F_{\mathcal{TL}(S_1)}\big(\Sigma,\mathcal{R}\big)$; and (iii) a symplectic function $\sigma(j) : \Sigma_{\sigma,S_1} \to \Sigma_{\sigma,S}$.

Let K be a (Borel) subset of the state space, $\Sigma_{\sigma,S}$; hence K represents a proposition about the system S. Then $\sigma(j)^*(K) := \sigma(j)^{-1}(K)$ is a subset of Σ_{σ,S_1} and, as such, represents a proposition about the system S_1. We say that $\sigma(j)^*(K)$ is the *pull-back* to Σ_{σ,S_1} of the S-proposition represented by K. The existence of such pull-backs is part of the consistency of the representation of propositions in classical mechanics, and it is important to understand what the analogue of this is in our topos scheme.

Consider the general case with the two systems S_1, S as above. Then let K be a proposition, represented as a sub-object of $\Sigma_{\phi,S}$, with a monic morphism $i_K : K \hookrightarrow \Sigma_{\phi,S}$. The question now is if the triple $\langle \tau_\phi(j), \phi(j), \beta_\phi(j)\rangle$ can be used to pull K back to give a proposition in $\tau(S_1)$, i.e., a sub-object of Σ_{ϕ,S_1}?

The first requirement is that the functor $\tau_\phi(j) : \tau_\phi(S) \to \tau_\phi(S_1)$ should *preserve monics*; for example by being left-exact. In this case, the monic morphism $i_K : K \hookrightarrow \Sigma_{\phi,S}$ in $\tau_\phi(S)$ is transformed to the monic morphism [41]

$$\tau_\phi(j)(i_K) : \tau_\phi(j)(K) \hookrightarrow \tau_\phi(j)(\Sigma_{\phi,S}) \qquad \text{in } \tau_\phi(S_1).$$

Thus $\tau_\phi(j)(K)$ is a sub-object of $\tau_\phi(j)(\Sigma_{\phi,S})$ in $\tau_\phi(S_1)$. It is a property of a topos that the pull-back of a monic morphism is monic ; i.e., if $M \hookrightarrow Y$ is monic, and if $\psi : X \to Y$, then $\psi^{-1}(M)$ is a sub-object of X. Therefore, in the case of interest, the monic morphism $\tau_\phi(j)(i_K) : \tau_\phi(j)(K) \hookrightarrow \tau_\phi(j)(\Sigma_{\phi,S})$ can be pulled back along $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)(\Sigma_{\phi,S})$ (see diagram (79)) to give the monic $\phi(j)^{-1}(\tau_\phi(j)(K)) \subseteq \Sigma_{\phi,S_1}$. This is a candidate for the pull-back of the proposition represented by the sub-object $K \subseteq \Sigma_{\phi,S}$.

Therefore, propositions can be pulled-back provided that the functor $\tau_\phi(j) : \tau_\phi(S) \to \tau_\phi(S_1)$ preserves monics.

7.2 The Topos Rules for General Crowd Dynamics

We will now present the general rules for using topos theory in the mathematical representation of crowd systems and their psycho-physical theories. The category $\mathcal{M}(\mathbf{Sys})$ is defined as follows:

1. The objects of $\mathcal{M}(\mathbf{Sys})$ are the topoi that are to be used in representing the systems in **Sys**.
2. The morphisms from τ_1 to τ_2 are defined to be the left-exact functors from τ_1 to τ_2.

The rules for using topos theory are as follows: [41]

1. A *topos realization*, ϕ, of **Sys** in $\mathcal{M}(\mathbf{Sys})$ is an assignment, to each system S in **Sys**, of a triple $\phi(S) = \langle \rho_{\phi,S}, \mathcal{TL}(S), \tau_\phi(S) \rangle$ where:
   (a) $\tau_\phi(S)$ is the topos in $\mathcal{M}(\mathbf{Sys})$ in which the theory-type applied to system S is to be realised.
   (b) $\mathcal{TL}(S)$ is the local language that is associated with S. This is independent of the realization, ϕ, of **Sys** in $\mathcal{M}(\mathbf{Sys})$.
   (c) $\rho_{\phi,S} : \mathcal{TL}(S) \rightsquigarrow \tau_\phi(S)$ is a representation of the local language $\mathcal{TL}(S)$ in the topos $\tau_\phi(S)$.
   (d) In addition, for each morphism $j : S_1 \to S$ in **Sys** there is a triple $\langle \tau_\phi(j), \phi(j), \beta_\phi(j) \rangle$ that interpolates between $\rho_{\phi,S} : \mathcal{TL}(S) \rightsquigarrow \tau_\phi(S)$ and $\rho_{\phi,S_1} : \mathcal{TL}(S_1) \rightsquigarrow \tau_\phi(S_1)$; for details see below.
2. (a) The representations, $\rho_{\phi,S}(\Sigma)$ and $\rho_{\phi,S}(\mathcal{R})$, of the ground symbols Σ and $\mathcal{R}$ in $\mathcal{TL}(S)$ are denoted $\Sigma_{\phi,S}$ and $\mathcal{R}_{\phi,S}$, respectively. They are known as the 'state object' and 'quantity-value object' in $\tau_\phi(S)$.
   (b) The representation by $\rho_{\phi,S}$ of each function symbol $A : \Sigma \to \mathcal{R}$ of the system S is a morphism, $\rho_{\phi,S}(A) : \Sigma_{\phi,S} \to \mathcal{R}_{\phi,S}$ in $\tau_\phi(S)$; we will usually denote this morphism as $A_{\phi,S} : \Sigma_{\phi,S} \to \mathcal{R}_{\phi,S}$.

   (c) Propositions about the system S are represented by sub-objects of $\Sigma_{\phi,S}$. These will typically be of the form $A^{-1}_{\phi,S}(\Xi)$, where Ξ is a sub-object of $\mathcal{R}_{\phi,S}$.
3. Generally, there are no 'microstates' for the system S; i.e., no global elements (morphisms $1 \to \Sigma_{\phi,S}$) of the state object $\Sigma_{\phi,S}$; or, if there are any, they may not be enough to determine $\Sigma_{\phi,S}$ as an object in $\tau_\phi(S)$. Instead, the role of a state is played by a 'truth sub-object' $\mathbb{T}$ of $P\Sigma_{\phi,S}$. In crowd mechanics, the truth object corresponding to a microstate s is the collection of all propositions that are true in the state s. If $\gamma \in \mathrm{Sub}(\Sigma_{\phi,S}) \simeq \Gamma(P\Sigma_{\phi,S})$, the 'truth of the proposition represented by γ' is defined to be [39]

$$\nu(\ulcorner\gamma\urcorner \in \mathbb{T}) = [\![\, \tilde{\gamma} \in \tilde{\mathbb{T}} \,]\!]_\phi \, \langle \gamma, \mathbb{T} \rangle.$$

4. There is a 'unit object' $1_{\mathcal{M}(\mathbf{Sys})}$ in $\mathcal{M}(\mathbf{Sys})$ such that if $1_{\mathbf{Sys}}$ denotes the trivial system in $\mathbf{Sys}$ then, for all topos realizations ϕ,

$$\tau_\phi(1_{\mathbf{Sys}}) = 1_{\mathcal{M}(\mathbf{Sys})}.$$

   Motivated by the results for quantum theory, we postulate that the unit object $1_{\mathcal{M}(\mathbf{Sys})}$ in $\mathcal{M}(\mathbf{Sys})$ is the category of sets:

$$1_{\mathcal{M}(\mathbf{Sys})} = \mathbf{Set}.$$

5. To each morphism $j : S_1 \to S$ in $\mathbf{Sys}$, we have the following:
   (a) There is a translation $\mathcal{TL}(j) : \mathcal{TL}(S) \to \mathcal{TL}(S_1)$. This is specified by a map between function symbols: $\mathcal{TL}(j) : F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R})$.
   (b) With the translation $\mathcal{TL}(j) : F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \to F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R})$ there is associated a corresponding function

$$\phi(\mathcal{TL}(j)) : \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}.$$

   These may, or may not, fit together in the commutative diagram [41]:

$$\begin{tikzcd}
F_{\mathcal{TL}(S)}(\Sigma, \mathcal{R}) \arrow[r, "\rho_{\phi,S}"] \arrow[d, "\mathcal{TL}(j)"'] & \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \arrow[d, "\phi(\mathcal{TL}(j))"] \\
F_{\mathcal{TL}(S_1)}(\Sigma, \mathcal{R}) \arrow[r, "\rho_{\phi,S_1}"'] & \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}
\end{tikzcd} \qquad (83)$$

   (c) The function $\phi(\mathcal{TL}(j)) : \mathtt{Mor}\tau_\phi(S)\Sigma_{\phi,S}\mathcal{R}_{\phi,S} \to \mathtt{Mor}\tau_\phi(S_1)\Sigma_{\phi,S_1}\mathcal{R}_{\phi,S_1}$ is built from the following ingredients. For each topos realization ϕ, there is a triple $\langle \tau_\phi(j), \phi(j), \beta_\phi(j) \rangle$ where:

i. $\tau_\phi(j) : \tau_\phi(S) \to \tau_\phi(S_1)$ is a left-exact functor; i.e., a morphism in $\mathcal{M}(\mathbf{Sys})$.
ii. $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)\big(\Sigma_{\phi,S}\big)$ is a morphism in $\tau_\phi(S_1)$.
iii. $\beta_\phi(j) : \tau_\phi(j)\big(\mathcal{R}_{\phi,S}\big) \to \mathcal{R}_{\phi,S_1}$ is a morphism in $\tau_\phi(S_1)$.

These fit together in the diagram [41]

$$\begin{tikzcd}
\Sigma_{\phi,S} \arrow[r, "A_{\phi,S}"] \arrow[d, "\tau_\phi(j)"'] & \mathcal{R}_{\phi,S} \arrow[d, "\tau_\phi(j)"] \\
\tau_\phi(j)(\Sigma_{\phi,S}) \arrow[r, "\tau_\phi(j)(A_{\phi,S})"'] & \tau_\phi(j)(\mathcal{R}_{\phi,S}) \arrow[d, "\beta_\phi(j)"] \\
\Sigma_{\phi,S_1} \arrow[u, "\phi(j)"] \arrow[r, "\phi(\mathcal{T}\mathcal{L}(j))(A_{\phi,S})"'] & \mathcal{R}_{\phi,S_1}
\end{tikzcd} \tag{84}$$

The morphisms $\phi(j)$ and $\beta_\phi(j)$ should behave appropriately under composition of morphisms in **Sys**.

The commutativity of the diagram (83) is equivalent to the relation

$$\phi(\mathcal{T}\mathcal{L}(j))(A_{\phi,S}) = [\mathcal{T}\mathcal{L}(j)(A)]_{\phi,S_1} \tag{85}$$

for all $A \in F_{\mathcal{T}\mathcal{L}(\phi,S)}\big(\Sigma, \mathcal{R}\big)$. As we keep emphasising, the satisfaction or otherwise of this relation will depend on the theory-type and, possibly, the representation ϕ.

(d) If a proposition in $\tau_\phi(S)$ is represented by the monic morphism, $K \hookrightarrow \Sigma_{\phi,S}$, the 'pull-back' of this proposition to $\tau_\phi(S_1)$ is defined to be $\phi(j)^{-1}\big(\tau_\phi(j)(K)\big) \subseteq \Sigma_{\phi,S_1}$.

6. (a) If S_1 is a sub-system of S, with an associated morphism $i : S_1 \to S$ in **Sys** then, in the diagram in (84), the morphism $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)(\Sigma_{\phi,S})$ is a monic morphism in $\tau_\phi(S_1)$.

In other words, Σ_{ϕ,S_1} is a sub-object of $\tau_\phi(j)(\Sigma_{\phi,S})$, which is denoted

$$\Sigma_{\phi,S_1} \subseteq \tau_\phi(j)(\Sigma_{\phi,S}). \tag{86}$$

We may also want to conjecture

$$\mathcal{R}_{\phi,S_1} \simeq \tau_\phi(j)\big(\mathcal{R}_{\phi,S}\big). \tag{87}$$

(b) Another possible conjecture is the following: if $j : S_1 \to S$ is an epic morphism in **Sys**, then, in the diagram in (84), the morphism $\phi(j) : \Sigma_{\phi,S_1} \to \tau_\phi(j)(\Sigma_{\phi,S})$ is an epic morphism in $\tau_\phi(S_1)$.

In particular, for the epic morphism $p_1 : S_1 \otimes S_2 \to S_1$, the morphism $\phi(p_1) : \Sigma_{\phi,S_1\otimes S_2} \to \tau_\phi\big(\Sigma_{\phi,S_1}\big)$ is an epic morphism in the topos $\tau_\phi(S_1 \otimes S_2)$.

One should not read Rule 2. above as implying that the choice of the state object and quantity-value object are *unique* for any give system S. These objects would at best be selected only up to isomorphism in the topos $\tau(S)$. Such morphisms in the topos $\tau(S)$[68] can be expected to play a key role in developing the topos analogue of the important idea of a *symmetry*, or *covariance* transformation of the theory (see [40]).

In the example of crowd mechanics, for all crowd systems we have $\tau(S) = \mathbf{Set}$ and $\Sigma_{\sigma,S}$ is a symplectic crowd manifold, and the collection of all symplectic manifolds is a general crowd category. It would be elegant if we could assert that, in general, for a given theory-type the possible crowd objects in a given topos τ form the objects of an *internal* crowd category in τ. However, to make such a statement would require a general theory of state objects and, at the moment, we do not have such a thing.

References

[1] Ivancevic, V., Ivancevic, T.: Human–Like Biomechanics. Springer, Heidelberg (2006)

[2] Ivancevic, V., Ivancevic, T.: Natural Biodynamics. World Scientific, Singapore (2006)

[3] Ivancevic, V., Ivancevic, T.: Geometrical Dynamics of Complex Systems. Springer, Dordrecht (2006)

[4] Ivancevic, V., Ivancevic, T.: Applied Differfential Geometry: A Modern Introduction. World Scientific, Singapore (2007)

[5] Freeman, W.J., Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Phys. Life Rev. 2(3), 93–118 (2006)

[6] Umezawa, H.: Advanced field theory: micro macro and thermal concepts. Am. Inst. Phys., New York (1993); Wiener, N.: Cybernetics. Wiley, New York (1961)

[7] Frölich, H., Kremer, F.: Coherent Excitations in Biological Systems. Springer, New York (1983)

[8] Penrose, R.: The Emperor's New Mind. Oxford Univ. Press, Oxford (1989)

[9] Penrose, R.: Shadows of the Mind. Oxford Univ. Press, Oxford (1994)

[10] Hameroff, S.R.: Quantum coherence in microtubules: A Neural basis for emergent consciousness. J. Consc. Studies 1(1), 91–118 (1994)

[11] Hameroff, S.R., Penrose, R.: Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. In: Hameroff, S.R., Kaszniak, A.W., Scott, A.C. (eds.) Toward a Science of Consciousness: the First Tucson Discussion and Debates, pp. 507–539. MIT Press, Cambridge (1996)

---

[68] Care is needed not to confuse morphisms in the topos $\tau(S)$ with morphisms in the category $\mathcal{M}(\mathbf{Sys})$ of topoi. A morphism from the object $\tau(S)$ to itself in the category $\mathcal{M}(\mathbf{Sys})$ is a left-exact morphism in the topos $\tau(S)$. However, not every morphism in $\tau(S)$ need arise in this way, and an important role can be expected to be played by morphisms of this type. A good example is when $\tau(S)$ is the category of sets, **Set**. Typically, $\tau_\phi(j) : \mathbf{Set} \to \mathbf{Set}$ is the identity, but there are many morphisms from an object O in **Set** to itself: they are just the functions from O to O.

[12] Hameroff, S.R.: Quantum computation in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness. Phil. Trans. Roy. Soc. London (A) 356, 1869–1896 (1998)

[13] Hagan, S., Hameroff, S.R., Tuszynski, J.A.: Quantum computation in brain microtubules: Decoherence and biological feasibility. Phys. Rev. D 65, 61901 (2002)

[14] Pribram, K.: Brain and the composition of conscious experience. J. Consc. Studies 6(5), 12–18 (1999)

[15] Pribram, K.: Holonomic brain theory. Scholarpedia 2(5), 2735 (2007)

[16] Ivancevic, V.G., Ivancevic, T.T.: Quantum Neural Computation. Springer, Heidelberg (2009)

[17] Ivancevic, V.G., Ivancevic, T.T.: Neuro-Fuzzy Associative Machinery for Comprehensive Brain and Cognition Modelling. Springer, Heidelberg (2007)

[18] Ivancevic, V.G., Reid, D.J., Aidman, E.V.: Crowd Behavior Dynamics: Entropic Path-Integral Model. Nonl. Dyn. 59, 351–373 (2010)

[19] Ivancevic, V.G., Reid, D.J.: Entropic Geometry of Crowd Dynamics. In: Evans, T. (ed.) Nonlinear Dynamics, Intech, Wiena (2010)

[20] Ivancevic, V.G., Reid, D.J.: Geometrical and Topological Duality in Crowd Dynamics. Int. J. Biomath. (in press)

[21] Ivancevic, V., Aidman, E.: Life-space foam: A medium for motivational and cognitive dynamics. Physica A 382, 616–630 (2007)

[22] Ivancevic, V., Aidman, E., Yen, L.: Extending Feynman's Formalisms for Modelling Human Joint Action Coordination. Int. J. Biomath. 2(1), 1–7 (2009)

[23] Ivancevic, V.G., Reid, D.J.: Dynamics of Confined Crowds Modelled using Entropic Stochastic Resonance and Quantum Neural Networks. Int. Def. Sup. Sys. 2(4), 269–289 (2009)

[24] Barr, M., Wells, C.: Toposes, Triples and Theories. Reprints in The. Appl. Cat. 1, 1–289 (2005)

[25] Illusie, L.: What is a Topos. Notices of AMS 51(9), 1060–1061 (2004)

[26] Lawvere, W., Schanuel, S.: Conceptual Mathematics: A First Introduction to Categories. Cambridge Univ. Press, Cambridge (1997)

[27] Lawvere, F.W.: Continuously variable sets: algebraic geometry = geometric logic. In: Proceedings Logic Colloquium Bristol 1973. North-Holland, Amsterdam (1975)

[28] Bell, J.L.: Toposes and Local Set Theories. Clarendon Press, Oxford (1988)

[29] MacLane, S., Moerdijk, I.: Sheaves in Geometry and Logic: A First Introduction to Topos Theory. Springer, London (1992)

[30] Lambek, J., Scott, P.J.: Introduction to higher order categorical logic. Cambridge University Press, Cambridge (1986)

[31] Intuitionistic Logic, Stanford Encyclopedia of Phylosophy. Stanford Univ., Stanford (2009)

[32] Balbes, R., Dwinger, P.: Distributive Lattices. Univ. Missouri Press, Columbia (1974)

[33] Borceux, F.: Handbook of Categorical Algebra 3. In: Encyclopedia of Mathematics and its Applications, vol. 53. Cambridge Univ. Press, Cambridge (1994)

[34] Isham, C.J.: Topos theory and consistent histories: The internal logic of the set of all consistent sets. Int. J. Theor. Phys. 36, 785–814 (1997)

[35] Isham, C.J.: A New Approach to Quantising Space-Time: I. Quantising on a General Category. Adv. Theor. Math. Phys. 7, 331–367 (2003)

[36] Isham, C.J.: A New Approach to Quantising Space-Time: II. Quantising on a Category of Sets. Adv. Theor. Math. Phys. 7, 807–829 (2004)

[37] Isham, C.J.: A New Approach to Quantising Space-Time: III. State Vectors as Functions on Arrows. Adv. Theor. Math. Phys. 8, 797–811 (2004)

[38] Döring, A., Isham, C.J.: A topos foundation for theories of physics: I. Formal Languages for Physics. J. Math. Phys. 49, 53515 (2008)

[39] Döring, A., Isham, C.J.: A topos foundation for theories of physics: II. Daseinisation and the liberation of quantum theory. J. Math. Phys. 49, 53516 (2008)

[40] Döring, A., Isham, C.J.: A topos foundation for theories of physics: III. Quantum theory and the representation of physical quantities with arrows $\breve{\delta}^{\circ}(\mathrm{A}): \underline{\Sigma} \to \underline{\mathbb{R}^{\succeq}}$. J. Math. Phys. 49, 53517 (2008)

[41] Döring, A., Isham, C.J.: A topos foundation for theories of physics: IV. Categories of systems. J. Math. Phys. 49, 53518 (2008)

[42] Döring, A.: Topos Theory and ‘Neo-Realist’ Quantum Theory. In: Fauser, B., et al. (eds.) Quantum Field Theory. Birkhauser Verlag, Basel (2009)

[43] Stone, M.H.: The theory of representations for Boolean algebras. Trans. Amer. Math. Soc. 40, 37–111 (1936)

[44] Mac Lane, S.: Natural associativity and commutativity. Rice University Studies, Papers in Mathematics 49, 28–46 (1963)

[45] Dosen, K., Petric, Z.: The Maximality of Cartesian Categories. Math. Logic Quart. 47, 137–144 (2001)

[46] Dosen, K., Petric, Z.: Proof-Theoretical Coherence. Studies in Logic 1. King’s College Publ., London (2004)

[47] Lambek, J., Scott, P.J.: Introduction to Higher-Order Categorical Logic. Cambridge Univ. Press, Cambridge (1986)

[48] Dosen, K., Petric, Z.: Bicartesiaan Coherence. Studia Logica 71, 331–353 (2002)

[49] Lambek, J.: Deductive systems and categories I: Syntactic calculus and residuated categories. Mathematical Systems Theory 2, 287–318 (1968)

[50] Lambek, J.: Deductive systems and categories II: Standard constructions and closed categories. In: Category Theory, Homology Theory and their Applications I. Lecture Notes in Mathematics, vol. 86, pp. 76–122. Springer, Berlin (1969)

[51] Lambek, J.: Deductive systems and categories III: Cartesian closed categories, intuitionist propositional calculus, and combinatory logic. In: Lawvere, F.W. (ed.) Toposes, Algebraic Geometry and Logic. Lecture Notes in Mathematics, vol. 274, pp. 57–82. Springer, Berlin (1972)

[52] MacLane, S.: Categories for the Working Mathematician. Springer, Heidelberg (1971)

[53] lang, S.: Algebra, 3rd edn. Springer, Heidelberg (2002)

[54] Goldblatt, R.: Topoi: The Categorial Analysis of Logic. North-Holland, London (1984)

[55] Kock, A.: Synthetic Differential Geometry. LMS lecture note series, vol. 51. Cambridge University Press, Cambridge (1981)

[56] Dalla Chiara, M.L., Giuntini, R.: Quantum logics. In: Gabbay, G., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. VI, pp. 129–228. Kluwer, Dordrecht (2002)

[57] Dowker, F.: Causal sets and the deep structure of space-time. arXiv.org/abs/gr-qc/0508109 (2005)

[58] Eilenberg, S., Mac Lane, S.: General theory of natural equivalences. Transactions of the American Mathematical Society 58, 231–294 (1945)

[59] Switzer, R.K.: Algebraic Topology – Homology and Homotopy. In: Classics in Mathematics. Springer, New York (1975)

[60] Stuart, J.: Calculus, 4th edn. Brooks/Cole Publ., Pacific Grove, CA (1999)

[61] Nash, C., Sen, S.: Topology and Geometry for Physicists. Academic Press, London (1983)

[62] Dieudonne, J.A.: A History of Algebraic and Differential Topology 1900–1960. Birkháuser, Basel (1988)

[63] Kan, D.M.: Adjoint Functors. Trans. Am. Math. Soc. 89, 294–329 (1958)

[64] Coecke, B.: Introducing categories to the practicing physicist. arXiv:physics.quant-ph.0808.1032 (2008)

[65] Coecke, B., Oliver, E.: Categories for the practising physicist. arXiv:physics.quant-ph. arXiv:0905.3010 (2009)

[66] Abramsky, S.: Categories, Proofs and Processes. Course at Oxford University Computing Laboratory (2002 ⟶ ...), Documentation and lecture notes `web.comlab.ox.ac.uk/oucl/courses/topics05-06/cpp/`

[67] Priest, G.: An Introduction to Non-Classical Logic, 2nd edn. Cambridge Univ. Press, Cambridge (2008)

[68] Mints, G.: A Short Introduction to Intuitionistic Logic. Springer, Heidelberg (2000)

[69] Fitting, M.: Proof Methods for Modal and Intuitionistic Logics. Springer, Heidelberg (1983)

[70] Gabbay, D.: Semantical Investigations in Heyting's Intuitionistic Logic. Springer, Heidelberg (2009)

[71] Miller, D.: Critical Rationalism. Open Court Pub. Comp., Chicago (2003)

[72] Popper, K.: Conjectures and Refutations: The Growth of Scientific Knowledge, 2nd edn. Routledge, New York (2002)

Chapter 4
Methodology for the Evaluation of an International Airport Automated Border Control Processing System

Veneta MacLeod and Brett McLindin

Defence Science and Technology Organisation (DSTO), Edinburgh, South Australia

Abstract. Biometrics are increasingly being used as a tool for identification and verification and are currently being implemented in access control situations, such as for border control. Biometrics are often used for such purposes on the assumption that it provides greater accuracy and security than humans performing these tasks and that there is the potential for greater efficiency in terms of processing times and resources. Nevertheless, the introduction of a biometric system, particularly where there are potential security implications, warrants considered evaluation before the system becomes operational. Preliminary evaluation may involve factory-acceptance testing, user-acceptance testing and scenario-based trials to determine likely operational performance. However, the most accurate assessment of system performance is obtained by an operational trial involving real travellers. This assessment should seek to determine: the operational performance of the biometric algorithm; how users (novice and experienced) interact with the system; and whether this interaction may impact on current and future business processes, as well as on the quality of the biometric samples obtained. This chapter presents a systems approach for evaluating traveller processing systems in the operational environment when implementing a new system or comparing an old with a new system. A system-level approach, that takes into account both technical performance and the impact of human factors issues, is recommended to provide a complete understanding of overall system performance, as well as identify potential improvements to enhance performance and/or useability.

1 Introduction

Automated border control systems using biometrics have recently been introduced in a number of airports around the world. These systems have been implemented in an attempt to increase security, combat fraud and improve traveller processing efficiency [1]. Automated border control systems are intended for use as a secure tool that performs some or all of the customs and immigration identity and processing tasks that are normally performed by airport officials at the Primary Line[1] for travellers arriving from international locations. Processing typically involves eligibility checks for entering the country and a facial comparison between the live traveller and the image from their travel documentation.

---

[1] The Primary Line is the first processing point for travellers entering or leaving a country. At this point the Primary Line Officer checks traveller identity and documentation and confirms the traveller's eligibility to enter or leave the country.

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 115–145.
springerlink.com

These measures have become prominent in anticipation of the influx of travellers expected following the introduction of A380 aircraft capable of transporting larger volumes of travellers[2]. Additionally, the 9/11 attacks in the US and the increasing ease with which individuals are able to travel the world have prompted the need for greater security in identifying travellers attempting to enter a country. These attacks prompted the US to initiate a requirement that countries wishing to continue to participate in a visa waiver program provide machine readable travel documents to their citizens [2, 3], although the earliest such passports were issued by Malaysia in 1998 [4]. To date, more than 60 countries have adopted electronic travel documents with this number expected to increase [4].

The specifications for these documents (typically called e-passports) have been defined by the International Civil Aviation Organization (ICAO). ICAO has specified that the e-passport has at least one biometric, with face to be included on all passports as a minimum. The facial biometric modality was chosen as the minimum specification as it is widely used, relatively easy to obtain in a non-intrusive manner, and may also be independently checked by a human [2]. In addition, an increasing number of countries are also incorporating fingerprints within their e-passports [2-3, 5]. The biometric and biographical information contained on the passport is stored on a chip, which is protected by a number of security measures [2].

As stated above, there is the assumption that the automated system that replaces the airport officials will improve security by more accurately identifying fraudulent travellers and will also improve traveller processing. However, before a system is fully implemented it is good practice to evaluate the automated system's performance to determine whether it improves, or at least maintains, existing performance measures. This type of assessment will quantify various performance metrics and determine the impact the introduction of the new system has on current (surrounding) business processes. This will input into the determination of the system's fitness for purpose (i.e., whether the automated system meets pre-defined specifications to determine whether it is meeting requirements, or whether the automated system would be appropriate as a (part) replacement tool of the manual Primary Line).

Defence Science and Technology Organisation (DSTO) has been conducting scenario- and operational-based evaluations of biometric systems since 2001. This experience has led to the development of a number of methodological protocols (in consultation with the Mansfield and Wayman's Best Practices paper [6]) for conducting both scenario and operational evaluations. A component-level evaluation (such as the independent evaluation of the technical performance of the biometric matching algorithm), is considered to be insufficient to obtain a complete understanding of system performance and is unlikely to provide an indication of operational performance in the airport. Instead, it is argued that a full system-level analysis that takes into consideration technical performance, as well as seeking to understand the influence of humans on the performance of the system should be conducted. This type of evaluation is considered imperative as it often identifies issues in using the system that would otherwise not be apparent (for example, a traveller having difficulty using the system because of unclear instructions).

---

[2] This increase in the number of travellers was expected to put a strain on the existing airport infrastructure.

In fact, previous DSTO evaluations of automated biometric systems in operational environments have found that the technical performance of such systems is quite high, but that it is the human aspects of the systems that limit overall system performance. However, often the impact of these factors is not quantified since many of these factors cannot be automatically logged by the system. For instance, if a large proportion of travellers are unable to insert their passports correctly because of hardware and/or human-systems integration issues, this will not be recorded but will prevent the travellers from actually reaching the biometric component of the system. This not only causes inconvenience for the travellers, but also fails to achieve the intended benefit of using the automated system to verify those travellers.

This chapter details a methodology for conducting an operational evaluation of an automated border control system within an airport environment. In order to obtain an appropriate system-level operational assessment of a system under consideration, it is suggested that performance should be considered in terms of true and false match rates, process timings (i.e., throughput) and qualitative observations. To evaluate differences in performance between the current manual Primary Line system and the new automated system (to determine whether there is a benefit of introducing the new system), these measures will need to be collected for both systems and then compared. In conducting operational evaluations of this nature, the common six-step process in conducting research should be followed:

1. **Definition of the problem space and a scoping of the analysis**, which involves determining the system boundary for both the automated border control system and the current manual system, so that the performance of the two systems may be evaluated and then compared;
2. **Determination of the quantitative and qualitative measures of performance** to determine which metrics will be used to define ‘performance’;
3. **Collection of the data**, which involves determining the best way to collect the defined metrics in an operational environment;
4. **Processing of the data**, which involves collating the raw data and putting it into a form so that it may be evaluated and then analysed;
5. **Analysis and interpretation of the data output**, which involves evaluating the manual and automated systems’ performance and comparing performance between the two systems. This data can then be used to assess the automated border control system’s fitness of purpose (i.e., whether it improves or maintains performance over the current system); and
6. **Reporting of the results**, which involves reporting the findings obtained from the trial, including a discussion of recommendations that may be made to the automated border control system or to the business processes to improve performance and functionality.

The specific application of each of these steps in the creation of a methodology for evaluating a particular automated border control system introduced into an airport environment is discussed in this chapter. It is hoped that the procedures outlined here may be adapted and be of use to other researchers conducting operational evaluations of other automated biometric systems.

2 Defining the Problem Space and Scoping the Analysis

The first step in conducting an assessment of this type is to determine what needs to be assessed (the problem space) and where the boundaries are. For example, it should be determined whether it is just the biometric matching algorithm that needs to be assessed, or whether it would it be useful to know how the overall system is performing.

Automated border control systems are integrated socio-technical systems that incorporate (hard) technical and (soft) human system components, with many of these aspects likely to influence performance. A component-level analysis of only the technical aspects of a biometric algorithm is unlikely to provide a complete understanding of the overall system level performance. Additionally, critical observations of the human/system interaction, which may lead to an understanding of how to improve performance of the system, may not be identified or their affect on system performance may not be quantified. It is thus argued that in almost all circumstances an overall system evaluation should be conducted to more effectively evaluate performance.

The introduction of an automated system into a new environment may pose questions regarding the matching accuracy of the system, its efficiency for processing travellers, whether travellers are able to use the system effectively, and level of traveller satisfaction and willingness to use the system again. Additionally, if the automated system is intended to replace (partially or completely) an existing system (in this case the manual Primary Line Operators), a comparative assessment between these two systems using the appropriate metrics should be conducted to determine whether the new system improves (or at least maintains) the level of performance over the previous system.

Considering a comparative assessment of the current system and the new automated system is proposed to obtain an assessment of performance, two system boundaries (one for each system) must be defined. A system and its boundary may be defined as:

> the area within which the decision-making process of the system has the power to make things happen, or prevent them from happening. More generally, a boundary is a distinction made by an observer which marks the difference between an entity he takes to be a system and its environment [7].

The primary aim of defining the system and its boundary before an evaluation commences is to identify potential variables that may affect the system under examination. The first step in this process is to define how each system is likely to operate.

The system boundaries for both the manual Primary Line and the automated border control system as we have defined it is presented in Figure 1 and described in the following sections.

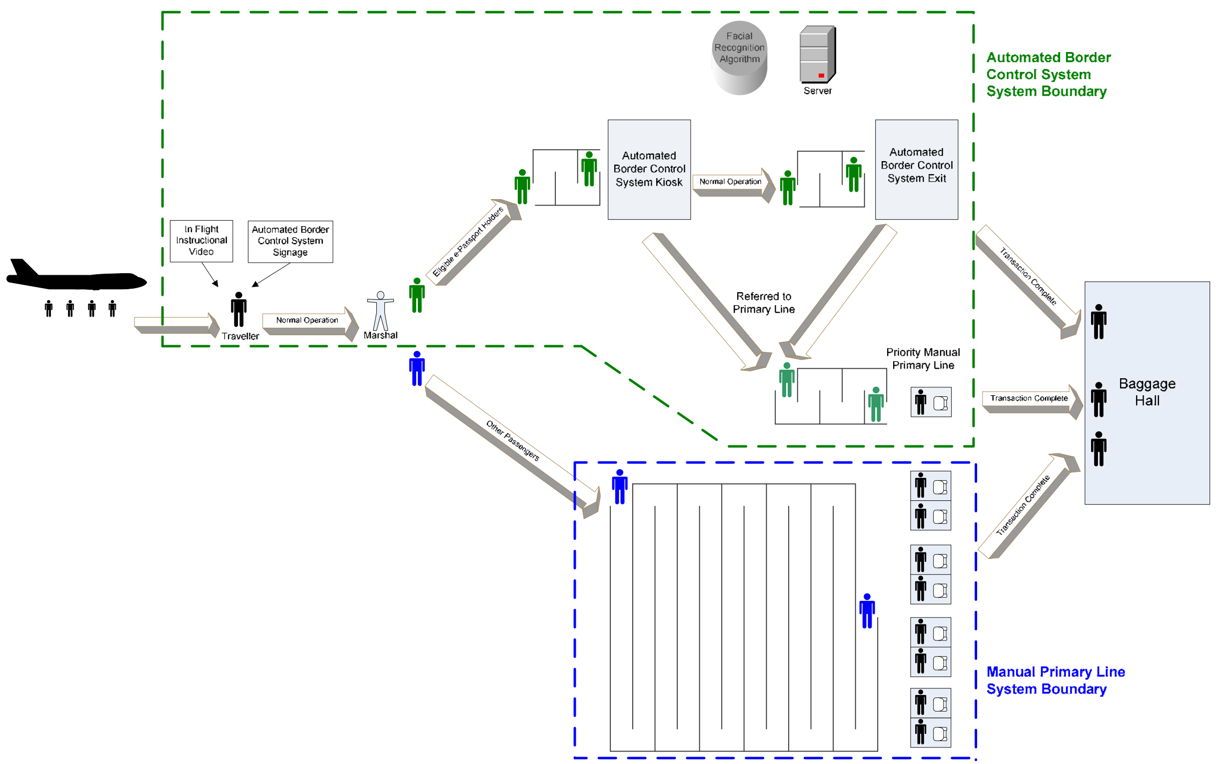


Fig. 1. Automated Border Control System and Manual Primary Line Border Control System Boundaries

2.1 Manual Primary Line Processing

Based on observations of systems in use and discussions with policy makers who have defined the current business practices, the procedures required for processing incoming international travellers could typically be described as follows:

1. Travellers arriving on international flights are given an Incoming Passenger Card (IPC[3]) to complete prior to or on arrival at the destination airport.
2. Once travellers have disembarked from their aircraft, they proceed to the arrivals area, with their hand luggage, for processing.
3. Travellers are directed to processing queues based upon their nationality and wait to be called to the Primary Line processing booth by a Primary Line Officer.
4. Once at the processing booth, the traveller must remove hats and sunglasses (if applicable) and present their IPC and passport to the Primary Line Officer for processing.
5. The Primary Line Officer then:
    a. scans and verifies the authenticity of the traveller's passport;
    b. confirms that the traveller is expected into the country;
    c. checks that the IPC has been completed;
    d. compares the face of the traveller with the documentation presented to determine whether they are of the same person; and
    e. asks any necessary questions of the traveller to confirm their identity and eligibility to enter the country.
6. If all of these criteria are met, the traveller is given back their IPC and passport and allowed to pass through into the arrivals hall to collect their checked in luggage. However, if the Primary Line Officer identifies an issue with the traveller's documentation or responses then they are referred either to Immigration or another Primary Line Officer for a more detailed inspection.

2.2 Automated Border Control System Processing

Based on observations of systems in use and discussions with system designers and integrators, traveller processing with an automated border control system could follow the following steps[4]:

---

[3] Travellers' answers on the IPC provide an indication to airport officials of what level of scrutiny should be applied to the traveller during customs and immigration processing.

[4] Note that this represents the procedure for an automated border control system where the initial processing and access control components of the system are separated. Other automated border control systems may possibly differ from this distributed system design and use other procedures.

1. Travellers arriving on international flights are given an IPC to complete prior to, or on arrival at the destination airport.
2. Once travellers have disembarked from their aircraft, they proceed to the arrivals area, with their hand luggage, for processing.
3. Travellers who are 18 years old or older and who hold an eligible e-passport are directed to the queues to use the automated border control system.
4. When a Kiosk is free, the traveller approaches the Kiosk, opens their passport to the photo page and places their passport into the passport reader face down.
5. The passport reader reads the passport and extracts the photographic image on the chip in the passport, creates a biometric template[5] and temporarily stores it in the system.
6. The traveller answers a series of questions to confirm their eligibility to enter the country and use the automated border control system.
7. If the traveller's eligibility to use the automated border control system is confirmed, a ticket is issued to the traveller. The traveller is then advised to remove their passport and continue to the access control portion of the system (the Exit) with their hand luggage.
8. If, for any reason, the traveller is unable to be processed using the automated border processing system they are advised to proceed to a particular manual Primary Line processing queue for priority processing. At this point, processing is similar to that for manual Primary Line processing (as discussed in Section 2.1).
9. At the Exit, the traveller is advised to insert their ticket into the ticket reader, remove their hat and sunglasses (if applicable) and look straight ahead at a set of flashing lights. While looking straight ahead, imagery is acquired of the traveller. The template created from that imagery is compared with the template previously created from the image on the e-passport.
10. If the algorithm of the automated border control system determines that the two templates match above a predetermined threshold, the ticket is returned and the Exit opens, allowing the traveller to proceed to collect their luggage and process through immigration.
11. If the imagery cannot be collected or the templates are deemed by the system not to match above the threshold, the ticket is retained and the traveller is referred to a particular manual Primary Line processing queue for priority processing. At this point processing is similar to that for manual Primary Line processing (as discussed in Section 2.1).

2.3 Defining the Components within the System Boundary

Of the two systems defined above, the system boundary for both systems needs to be identified and the components within that boundary examined for the evaluation.

---

[5] A template is defined as the stored reference measure of a user generated from features extracted from enrolment samples [6].

2.3.1 Manual Border Control System Boundary

The major components identified as being within the manual Primary Line system boundary (i.e., those components that can have a direct impact on performance of the system) include the:

- Primary Line processing queues;
- manual Primary Line processing booth;
- Primary Line officer; and
- travellers.

These components are described in detail in the following sections.

2.3.1.1 Primary Line Processing Queues

The manual Primary Line processing queue is where the travellers stand and wait prior to processing.

2.3.1.2 Manual Primary Line Processing Booth

The manual Primary Line processing booth is the location where the traveller identity checks are performed by the Primary Line Officer. It incorporates a computer, passport scanner, and other traveller verification systems.

2.3.1.3 Primary Line Officers

Primary Line Officers currently perform several tasks when assessing travellers. These include tasks that are able to be performed by the automated border control system, such as the expected movement check to determine whether the traveller was scheduled to arrive at that destination at that time, and the facial recognition comparison of a passport image to the presenting traveller. However, over and above the assessments that are able to be made by automated systems, human operators are able to perform other security checks, such as an evaluation of the traveller's behaviour, which can help guide assessments on authenticity.

2.3.1.4 Travellers

The traveller constitutes any individual who is processed at the manual Primary Line processing booth. If direct comparisons are to be made between the manual Primary Line system and the automated border control system, similar populations of travellers should be assessed for each system (e.g., if the automated border control system is only able to be used by individuals 18 years and older, than the comparative travellers in the Primary Line should also be limited to individuals 18 years and older).

2.3.2 Automated Border Control System Boundary

The major hard and soft system components of this particular system that were identified as being within the automated border processing system boundary include the:

- processing queues for the Kiosk and the Exit;
- Kiosk;
- Exit;
- facial recognition matching algorithm;
- areas between the Kiosk and the Exit (as it is a distributed system);
- priority manual processing booth (if transaction at the automated system is unable to be completed);
- marshals who guide the travellers where applicable; and
- travellers.

2.3.2.1 Automated Border Control System Processing Queues

The automated border access control system processing queues are where the traveller waits for the next available Kiosk or Exit to be available before processing.

2.3.2.2 Kiosk

In this automated border control system the Kiosk comprises:

- a passport reader;
- a token printer;
- a user interface screen; and
- connections to a server and other Customs systems.

The Kiosk initiates and returns an assessment on a traveller's eligibility for automated processing. Travellers holding an eligible passport open their passport to the photo page on the passport and insert it into the passport reader to commence their eligibility check for processing. The eligibility check involves retrieving the traveller's facial image and personal data that is stored on the chip of the passport and comparing it to passenger expected movement details to determine if the traveller was scheduled to be arriving at that particular destination. The retrieved information is stored on the system in preparation for the expected facial recognition task at the Exit. Travellers are then asked a series of questions via the user interface screen (such as their contact with particular diseases) to determine their eligibility to enter the country without further investigation. If the traveller is able to continue to self-process using the automated system, a template of the traveller's passport image is created and stored for use at the Exit. The traveller is then issued with a ticket that will be used at the Exit to call up their record.

If the traveller is unable to be processed using the automated border control system, then they are referred to the priority manual processing booth (see Section 2.3.2.6) for manual processing.

2.3.2.3 Exit

In this automated border control system the Exit comprises:

- a gate;
- a ticket reader;
- two lighting towers;

- a totem (containing three cameras and a user prompt screen);
- a facial recognition algorithm; and
- connections to a server and other Customs systems.

After travellers have successfully completed the assessment eligibility process at the Kiosk they are directed to proceed to the Exit to complete processing. At the Exit, the traveller is required to insert the ticket that was issued at the Kiosk into the Exit ticket reader. Once the ticket has been inserted, the Exit validates the ticket and performs a traveller identity check. This involves imagery being acquired from the three cameras situated in vertical alignment on the totem. Imagery is collected from all three cameras simultaneously until the algorithm determines which camera is acquiring the best facial imagery, based on the traveller's height. When this decision has been made lights flash around the chosen camera to focus the traveller's attention to that particular camera for better quality imagery acquisition.

The obtained imagery is combined into a composite template, that is then compared with the template created from the traveller's passport image to determine whether they match. If a match is made above a pre-determined threshold, the traveller is deemed to be verified and the traveller's ticket is returned. The traveller is instructed to remove the ticket from the reader, causing the Exit to open and allowing the traveller to proceed to the baggage collection area. If a match is not achieved, the traveller's ticket is retained and the traveller is referred to the priority manual processing booth (see Section 2.3.2.6) for manual processing[6].

2.3.2.4 Facial Recognition Matching Algorithm

The facial recognition procedure commences at the Kiosk when the traveller inserts their e-passport into the passport reader. This triggers the system to open and then read the image stored on the chip of the e-passport. This image is then converted into a template which is stored until the traveller commences transacting at the Exit.

At the Exit, an additional template is created from the imagery of the traveller and is matched against the template previously created from the passport. If a match above a predetermined threshold is achieved, the traveller is permitted to enter the baggage hall. If the match falls below the threshold, the traveller is referred for manual processing.

2.3.2.5 Area between the Automated Border Control System Kiosk and Exit

As the Kiosk and the Exit components of this automated border control system are separated, travellers must walk between the two areas in order to continue processing. Travellers cross this space to use the Exit of the automated border control system if a ticket has been issued at the Kiosk authorising them to continue automated processing, or to reach the priority manual processing booth if unable to use the automated system.

---

[6] Note that if the automated border control system under examination is not distributed, that is, it does not have separate Kiosks and Exits, the ticket process described above is negated.

2.3.2.6 Priority Manual Processing Booth

Travellers who have attempted to transact with the automated border control system, but who were unable to be processed at either the Kiosk or the Exit, for whatever reason, are referred to a priority manual processing booth to complete their transaction. Processing at this point is similar to that which is conducted at the manual Primary Line (see Section 2.1).

2.3.2.7 Marshals

Although the aim of the automated border control system is for travellers to process independently, marshals may be used to instruct travellers on how to use the system or to answer any questions.

2.3.2.8 Travellers

A traveller constitutes any individual who is processed by the automated border control system. Eligible travellers (currently individuals aged 18 years and over with an e-passport) will have the option of using the automated border control system or to be processed manually at the Primary Line.

3 Determining the Quantitative and Qualitative Measures of Performance

Once the system boundary has been identified, the quantitative and qualitative measures of performance for both systems are developed. Comparing both systems using the same measures of performance enables the determination of whether the automated border control system confers an advantage over, or at least performs to the same level as, the existing manual Primary Line processing. In particular, this type of assessment is needed to guide the evaluation of whether the automated border control system would be appropriate as a (part) replacement tool of the manual Primary Line. An understanding of the level of performance of both systems on these defined measures may identify issues that could be modified to improve performance and functionality.

Thus, as described in the following sections, it is recommended that a system level analysis of the automated border control and manual Primary Line systems encompass an assessment of both technical and human factors aspects[7], including an assessment of the following key areas:

- technical performance of the physical components of the system;
- matching performance;
- process timings of users of the system;

---

[7] Human factors is the study of human systems within the workplace in order to understand human work processes and their strengths, abilities, limitations, trends and interactions. The aim of this type of investigation is to develop optimal relationships between humans and the systems within which they operate, in order to maximise work efficiency [8].

- observations of the interaction between the users and the system; and
- travellers' perceptions of the system.

3.1 Technical Performance

The primary measures of performance for the technical aspects of the automated border control system can incorporate the systems':

- hardware components;
- software components; and
- combined information technology (IT) components.

This chapter does not focus on the quantification of these components specifically, as other evaluation methodologies of these components are widely available.

3.2 Matching Performance

Matching performance determines the expected level of accuracy in matching decisions that may be achieved by the system.

In relation to the biometric algorithm of the automated border control system this may be quantified in terms of the False Match Rate (FMR) and the False Non-Match Rate (FNMR). The FMR shows the proportion of travellers that are incorrectly verified and granted access by the system. This could have potential security implications by allowing fraudulent travellers to enter a country. The FNMR shows the proportion of travellers that are incorrectly rejected as not matching their passport. That is, the traveller has presented a valid passport of themselves, but the facial recognition algorithm has determined that the image from the passport and the live imagery collected of the traveller do not match. A false non-match could occur for several reasons, including:

- the imagery collected during the traveller's presentation at the Exit contained multiple travellers;
- the traveller was not looking at the camera;
- the traveller was not in the field of view during imaging (perhaps attending to hand luggage, etc);
- the algorithm cannot match the traveller above the threshold;
- the algorithm threshold is set too high; and/or
- the traveller no longer looks like the picture in their passport.

A false non-match result is a disruption for the traveller as the traveller is referred for further processing. This increases the processing time for falsely rejected travellers, and consumes additional (manual) resources. FMR and FNMR assessments may be undertaken for the sample as a whole or for different demographic groups (such as gender and ethnicity) to determine relative performance for various sectors of the population.

Human operator matching performance may also be assessed using FMR and FNMR metrics, as well as the measures of discrimination and bias that are derived from Signal Detection Theory. Discrimination indicates the ability of an individual to correctly distinguish a target from a non-target (i.e., an impostor image from a true match image), whereas bias indicates the propensity for an individual to raise an alarm when they believe a threat (i.e., a non-matching image) is present. These measures allow for a more comprehensive description of human performance on tasks of this nature than just detection accuracy statistics alone. Further details on these measures are detailed in [9].

3.3 Process Timings

It is reasonable to assume that one of the considerations of implementing an automated border control system is to reduce (or at least maintain) traveller processing times. The collection of process timings would also provide a range of times that may be expected for travellers using each system, which would determine whether each system is meeting predicted (and expected) timing performance. The analysis of process timing data would allow a comparative assessment between the two systems to determine if there are statistically significant and practical differences in timings between them[8]. This would provide some indication of how the implementation of the new automated system will impact (either positively or negatively) on business processes and traveller throughput. Additionally, this assessment may highlight resource issues with either system, such as the adequacy of Kiosk and Exit and Primary Line processing booth numbers within the airport for current (and future) traveller load.

It is also possible that demographics and other traveller variables may have an impact on processing times. Consequently, data on variables that may influence either the manual or automated system process timings (e.g., gender, approximate age, whether travellers wear glasses or hats or not, and whether travellers attempt to process singly or in groups) should also be collected in conjunction with timings. The identification of variables that significantly increase process timings may lead to modifications to the system to reduce this impact.

Additionally, experience at using the automated border control system may also influence process timings as it is likely that travellers will become more familiar with the use of the system over time and may theoretically process faster with subsequent usage. Process timings for novice and experienced users should therefore be obtained and compared. Such results would more closely define the likely range of times that could be expected from users of the automated system, with average process timings from experienced and novice users defining the minimum and maximum times that could be expected.

In order to undertake a comparative assessment of process timings, traveller processing times for both systems would need to be measured. Given that the automated border control system examined here consists of two processing components (the Kiosk and the Exit) and travellers must interact with both parts, processing times for the automated system can only be made if both aspects of the system are assessed. Additionally, as the time that travellers spend in other parts of the system (such as

---

[8] Statistical significance refers to the probability that obtaining a statistic of a given size due strictly to random sampling error or chance is less than the selected alpha size. Practical significance, however, refers to a value judgement about whether the statistic has real world relevance [10].

waiting in traveller queues or walking between the Kiosk and the Exit) also constitutes a part of the overall process, these components should also be assessed.

Thus, to adequately assess both systems, process timings should be collected for the following sub-processing components:

- queuing time for the automated border control system Kiosk;
- interaction time with the automated border control system Kiosk;
- movement time between the automated border control system Kiosk and the Exit;
- queuing time for the automated border control Exit;
- interaction time with the automated border control Exit;
- queuing time for the automated border control system priority manual processing booth;
- interaction time with the automated border control system priority manual processing booth;
- queuing time for the manual Primary Line processing booth; and
- interaction time with the manual Primary Line processing booth.

Each traveller should be tracked through each part of the system (whether the automated border control system or the manual Primary Line system) in order to obtain accurate overall process timings for a single traveller to process through that entire system.

3.4 Observations

The introduction of an automated biometric system may be intended to improve functionality, reduce processing times, and/or improve security. However, the intended benefit is lost if users of that system are unable to process effectively. Thus, human factors observations of the travellers transacting with the automated border control system, in conjunction with the collection of process timings (as discussed in Section 3.3), are required to assess overall performance. An observational assessment would identify issues within the automated system that prevent or hamper usage with the goal of improving the system by remedying these problems. This information is important as travellers will be directed to the manual processing line if they have difficulties using the system. Extra travellers at the Primary Line will impact on the Primary Line resources, as well as potentially increasing traveller process times and frustrations of the traveller.

An observational assessment may identify issues that result in:

- Difficulty in using the automated border control system because of usability issues (such as ergonomic issues, inadequate or unclear instructions, and the size, colour, placement and amount of text and buttons on screens, etc.) that were not apparent or considered during system design.
- Human/system interaction issues (e.g., travellers not understanding how to insert their passport into the passport reader, not realising that they were required to take their returned ticket to make the Exit open, or not knowing what to do if their ticket was not returned).

- The acquisition of poor quality biometric samples, which may reduce the ability of the algorithm to successfully accept or reject matches. For example, a poorer matching template will be acquired if travellers are not aware that they should be looking at the camera during biometric sample acquisition.
- Increased or decreased process timings as a result of particular behavioural characteristics of the travellers. For example, a group of travellers travelling and processing at the automated border control system together may (not realising that each ticket issued at the Kiosk is unique to each traveller) combine the tickets between the Kiosk and the Exit and therefore not transact with the correct ticket at the Exit and thus be rejected by the system. Additionally, processing at the Primary Line may be increased if travellers talk to the Primary Line Officers when at the processing booth.

In addition to identifying the specific factors that either prevent travellers processing or increase travellers' processing times, descriptive statistics on the number and types of incidences that resulted in such difficulties should also be collected. A better understanding of what factors cause the most transaction failures and which have resulted in marked differences in process timings may be useful to highlight where efforts to modify the system to maximise system efficiency should be directed.

3.5 Traveller's Perceptions

A traveller's perception (before and after) using an automated biometric system may guide their willingness to use the system subsequently. These perceptions may be influenced by beliefs on how usable the system is, consideration of the safety or security implications of using the system, and general thoughts on the use of technologies in preference to interacting with a human. These beliefs may be guided by previous experience with using this (or some other) automated system, or influenced by exposure to media, signage, information on the system (e.g., in-flight videos, traveller assist personnel), or the views of others (e.g., family, friends or co-workers). If a traveller's perception or experience with using the automated system is negative, they may be unwilling to use the system in the future, which may impact upon future business processes within the airport. If the reasons travellers (who either have or have not used the automated system) are reluctant to use the automated system when processing at airports in the future can be identified, these aspects could potentially be modified to improve travellers' perceptions.

Travellers' perceptions could be assessed by asking them to complete a questionnaire (or participate in a brief interview). Perceptions could either be collected both before and after travellers have been processed by the system (which would enable responses to be compared to determine whether exposure to the system resulted in changed perception), or only after travellers have been processed (to get an overall view of traveller perceptions). Responses could be collected in a free-text format, a 'yes' or 'no' response, or on the basis of a Likert scale rating (e.g., 1 = "Strongly Agree" to 5 = "Strongly Disagree"). This questionnaire/interview could seek to determine (amongst other variables of interest):

- demographic details of travellers using the systems;
- how the traveller would rate their experience of using either the manual or the automated border control system;
- the factors that prompted the traveller to use the automated border control system over the manual Primary Line system (or vice versa);
- whether they had heard about the use of the automated border control system from other people, media advertising or airport signage;
- whether (and how) their perception of using the automated border control system was positively or negatively influenced by the information that they had received on it from other people, media advertising or airport signage;
- whether (and how) their pre-exposure perceptions on the automated border control system from other people, media advertising or airport signage had changed after using the automated border control system for themselves;
- whether (and how and why) the information that the traveller had received on the automated border control system from other people, media advertising or airport signage was helpful or not in guiding their use of the automated border control system;
- how they would rate the adequacy of instructions that were provided by the automated border control system;
- which feature(s) of the automated border control system the traveller had difficulties with (if any) and whether they required assistance (and what type) to use the system;
- which feature(s) of the automated border control system the traveller did and did not like;
- whether, if given the chance, the traveller would choose to use the automated border control system again; and
- whether the traveller would recommend using the automated border control system to other people.

In order to examine a wide range of respondents, travellers who were successful (i.e., attempted and completed a transaction), as well as those that were unsuccessful (i.e., attempted to transact, but were referred, for whatever reason, to the priority manual processing booth) at using the automated system should be sampled. Additionally, eligible travellers who chose not to use the automated system should be sampled to determine their perceptions and reasons for choosing not to use the automated system. In conjunction with the qualitative observations described in Section 3.4, these insights may suggest common problems with the automated border control system, and suggest ways in which the system could be modified to improve functionality and higher traveller acceptance. Additionally, traveller's perceptions on the manual Primary Line and the automated border control system could be compared to determine whether their perceptions on the automated system differed according to whether they did or did not use this system to transact.

This analysis could also attempt to determine the impact of different types of signage (in terms of ability to attract people's attention) by using eye tracking technology to assess the percentage and length of time that travellers look at each sign. The presence of situational awareness cameras recording individuals processing through each component of the system (i.e., automated border control system signage, Kiosk, Exit and response at questionnaire) may further enable the correlation of eye tracking data with traveller processing times and responses on questionnaires, which will enhance understanding of how these factors relate to each other.

4 Collecting and Processing the Data

Following the identification of the quantitative and qualitative measures of performance that should be collected during the evaluation, the data collection techniques that are required to obtain this data must be defined. An assessment of the systems should be continued for as long as it takes to obtain a sample large enough to conduct statistical analysis on. For a starting point to obtain sample size calculations see [11].

The processes and equipment discussed in the following section have been identified as suitable to collect the data necessary for evaluating the measures of performance that were defined in Section 3.

4.1 Ground Truth

Ground truth data relating to the automated border control system is required to keep a record of what occurred during the operational evaluation. It is recommended that environmental variables such as temperature, relative humidity and ambient light levels be recorded using data loggers. This information may be required later to explain observed system performance variations.

Additionally, if the automated border control system provides it, measures of performance for image quality assessment for either the passport image or the individual verification images should be obtained and reported.

4.2 Matching Performance

During traveller transactions the automated border control system retrieves the image from the passport that has been inserted into the Kiosk by the traveller, stores it, and then converts it into an (enrolment) template. The traveller's passport image and the enrolment template, as well as the imagery and template obtained following traveller interaction at the Exit, should be stored for subsequent analysis of the matching performance of the algorithm. The outputs of the algorithm generated from the matching process are saved and analysed to determine performance in terms of FMR and FNMR. Cumulative probability plots are also generated.

System logs should be collected periodically throughout the operational evaluation to perform FNMR estimates to ascertain that the system is performing as expected. At the end of the evaluation the final FNMR statistics can be calculated. Additionally, if any changes are made to the system during the operational evaluation these changes should be recorded and later evaluated to determine whether they had an affect on the results. This assessment may determine whether the data can be analysed as a whole at the end of the evaluation (because the change was unlikely to have affected the data obtained) or whether the data should be segmented before and after the change and evaluated separately.

Operational evaluations offer the best data in terms of operational applicability. However, operational evaluations may not be practical to quantify the FMR metric using online transactions. That is because even if travellers fraudulently using the system were detected, there would not be enough of these cases to obtain a statistically significant sample from which to draw conclusions. Thus, to estimate these metrics with the automated border control system the traveller transactions are processed offline. In this process, every traveller's passport image (template) is compared against every other travellers' transaction template. This is classed as a Zero Effort level of attack (as defined in [12]) and would result if a traveller, after finding a passport, attempted to transact with that passport regardless of whether the gender, ethnicity and general appearance of the image on the passport matched their own or not. A more determined (fraudulent) traveller would likely try to obtain a passport with similar features to their own (gender, ethnicity, etc.).

FMR, FNMR, discrimination and bias metrics within the manual Primary Line system could also be evaluated by a scenario-based trial where individuals are required to present photographic identification to a human operator, who must determine whether the presenting individual matches the individual in the photograph. To obtain matching performance statistics a small percentage of presenting individuals would present false photographic identification (impostor images). Alternatively, such an assessment could be conducted in an operational environment where a small pool of individuals with impostor identification would join a stream of (one assumes) genuine travellers and attempt to by-pass the manual Primary Line with their false documentation. In both instances, human operator responses in terms of the number of travellers presented to each operator and their responses (correct match, incorrect match, correct reject and incorrect reject) for each presented traveller should be recorded (see [13]). These responses could then be used to calculate the FMR, FNMR, discrimination and bias metrics according to the equations presented in [9].

4.3 Process Timings and Observations

Process timings could be made on the basis of system logs, such as defining the total Kiosk processing time as being the time from when the passport reader registers that a passport has been inserted until the passport has been removed. However, it is argued that this way of assessing processing times does not effectively take into account the total time that a traveller takes to actually use the system. For instance, a traveller may not understand how to correctly insert their passport into the passport reader (e.g., which page they need to open the passport to, which way up the passport should go into the passport reader, and how long they need to keep the passport in the passport reader etc.) and may spend some time trying to accomplish this task. The time taken for travellers to understand how to use the system is representative of processing time, as it influences how long that traveller has to spend interacting with the system (and may potentially impact throughput if other travellers are prevented from using the system because of processing delays). However, these timings would not be recorded by system logs that only register timings once a passport has been successfully read. A more suitable means of collecting process timings is through manual observation of users of the system, to take into account unexpected human/system interaction issues.

Process timings and observations may thus be collected by:

- real-time observations; and/or
- video (and perhaps audio) recording during the evaluation.

There are advantages and disadvantages to both modes of data collection, which are outlined in Table 1 below.

Table 1. Advantages and Disadvantages of Data Collection via Real Time Observations and Video Recording

Real-Time Observations		*Video Recording*	
Advantages	*Disadvantages*	*Advantages*	*Disadvantages*
• reduces the risk of losing data or no data being collected if there is a problem with the camera/recording equipment; and • if there is a large area to be examined it may be more practical than having cameras record the whole area.	• requires a number of individuals onsite to collect data; • because of the manually intensive and time consuming data collection process, fewer samples are likely to be collected; • there is no option to recheck data if the start/stop times are missed, so the accuracy of the data cannot be verified; • the presence of people onsite observing may distract travellers and may cause them to change their behaviour (e.g., ask the observers what they are doing or ask the observers questions about the automated border control system procedure etc.) which may corrupt some of the data; and • only a limited amount of data can be observed (generally only one aspect can be observed at a time) and therefore important key elements may be lost.	• video footage can be replayed and rechecked if required to maintain accuracy; • time date stamped video footage enables more accurate timings to be collected; • data collection can be conducted off site at a more convenient time; • data collection can be conducted by multiple people in parallel, thus speeding up the analysis and reporting and also enabling more samples to potentially be collected; • this type of data collection is generally unobtrusive and is thus likely not to impact on traveller behaviour; • common observations or problems with the system can be shown to the stakeholders initiating the assessment to better illustrate aspects of the system where modifications might be recommended; and • there is a permanent record of each transaction in case there are any fraudulent attempts.	• there needs to be compliance from the stakeholders to set up such equipment; • there are costs involved in leasing/ buying and setting up the equipment; • staff need to be willing to monitor and change the recording media and to report any problems; • if there is a problem with the camera/ recording equipment, data may be lost; and • situational awareness may be lost if video cameras are not able to observe the entire area.

Obtaining video (and audio) footage is useful, if not critical, in conducting evaluations of this type. Nevertheless, it is recognised that there are often constraints imposed on the installation of recording equipment into certain operational environments, which may preclude using this method. In these instances, real-time manual observations will have to be used.

Regardless of the data collection method employed, it is recommended that prior to the assessment a number of on-site observations be conducted to obtain a proper understanding of the functioning of the system. Once it is understood how travellers use each of the systems, the anticipated start and stop times (and the associated cues for each) for the overall process times and for each individual subcomponent (as was identified in Section 3.3) can be defined. Additionally, on the basis of on-site observations the most common variables that impact travellers use of each system will be identified. This could then be used to generate a 'tick and flick' spreadsheet for data collection to determine the frequency of each type of error experienced during the operational trial. Examples of observations that may be relevant include when the traveller:

- is uncertain as to which page their passport should be opened to for transacting;
- removes their passport from the passport reader too quickly;
- fails to proceed to the Exit after obtaining a ticket at the Kiosk;
- is unaware that photographs will be taken of them at the Exit and thus fails to look straight ahead at the cameras for image acquisition; or
- is unaware that their ticket must be collected at the Exit after transaction to activate the gate so that they may proceed into the baggage hall.

Irrespective of the data collection method, data collection of this nature is ideally conducted by one individual (or only a small number of individuals). While there is some advantage in having multiple people collecting data, as more data may be collected and there may be better tracking of travellers, there is less consistency in application with different observers perhaps making minor (and usually unintentional) modifications to the data collection and therefore the results.

If real-time manual observations are collected, all observers will need to be able to determine the defined start and stop times (and associated cues) for all relevant components and be able to identify the most common observations in order to accurately and consistently record the required data. Each observer should be issued with paper copies of the identified observational categories, from which they would tick any categories that occur during the transaction for each traveller they observe processing. In addition, basic demographic details that may affect processing (such as gender, approximate age, whether traveller is wearing hats/glasses, whether traveller is travelling in a group) should be recorded[9]. Each observer

---

[9] It should be noted that if the assessment of differences between novice and experienced travellers is of interest, this data is unlikely to be easily identifiable by simple observations of users in an operational environment. This type of assessment may thus be more suited for evaluation in a scenario based trial where level of experience can be more readily controlled.

should also have a stop watch to record processing time in conjunction with the observations and demographic details collected for each traveller. Observers should attempt to remain as unobtrusive as possible to reduce both disruption to the operational environment and to avoid influencing the responses of the travellers.

If video footage is able to be collected, traveller transactions will need to be recorded from sufficient camera sources and angles to enable the identification of the most common errors travellers experience and to calculate process timings for both the automated border control and the manual Primary Line systems. Possible obstructions to the field of view of the camera from people using the system should be considered when defining camera angles that are appropriate for collecting the required data. The video footage should be time/date stamped to <1 second resolution so that footage from several different camera angles (if necessary) can be synchronised and viewed simultaneously when individual process times and observations are collected. Video equipment should be tested and calibrated before the trial to ensure that that the camera angles are appropriate for collecting the required data and that the equipment is functioning as expected. During the trial, enough media will need to be provided to record video footage for each camera for the entire duration. Staff will also need to be aware and willing to change the media and continue the recording at the required times.

To assist with the subsequent analysis, process timings and observational data should be recorded in an electronic spreadsheet. Process timings obtained from video recordings may be collected in one of two ways. Firstly, the video footage could be viewed and the start and stop times of each process on the basis of the time date stamp on the footage could be entered into the spreadsheet. Alternatively, a computer macro could be developed, where pressing a certain button (or combination of buttons) would enter the real time of the computer being used into the spreadsheet. Previous experience collecting timing data using both methods suggests that the use of the time/date stamp to note times is more useful for long processes (measured in minutes or hours) where (to save time in inputting the data) video footage may be fast-forwarded to the relevant stop time. However, for shorter processes (seconds) it is more useful to create a timing macro using a keyboard shortcut to record start and stop times in real time.

To adequately assess and compare process timings and observations of travellers' interaction issues between the two systems, data would need to be obtained by observing travellers at each stage of the process, namely the:

- automated border control system:
    - Kiosk processing queue;
    - Kiosk;
    - Exit processing queue;
    - Exit;
    - priority manual processing booth queue; and
    - priority manual processing booth;

- area between the automated border control system Kiosk and Exit;
- manual Primary Line processing queue; and
- manual Primary Line booth[10].

Each of these processes may also be furthered examined according to their individual sub-processes. For example, processing at the Kiosk could be divided into the following sub-components:

- from when each traveller first attempted to open their passport to the correct page;
- to when the traveller correctly inserts their passport into the Kiosk's passport reader;
- to when the traveller has finished answering the Kiosk questions and a ticket is printed; and
- until the traveller has finished processing and has gathered their bags/possessions and left the Kiosk area ready for another traveller to commence processing.

Assessing processing times to this level of detail may identify particular issues with each sub-component time, such as difficulties understanding instructions or problems with the passport reader. Overall Kiosk processing time would be from the time the traveller began opening their passport to the correct page to when they have collected their possessions and moved away from the Kiosk. Video cameras recording the Kiosk area should ensure that the passport reader and the ticket reader are visible and are not obscured by a traveller transacting. This same process should be conducted for each individual component of the system.

To obtain the total transaction time for a single traveller to be processed by the system, each traveller should be tracked as they progress through the entire system (i.e., the time from the traveller entering a queue until processing has been completed and the traveller enters the baggage hall). Tracking travellers through the entire system will require a number of situational cameras to observe all components of the system with which travellers may interact.

4.4 Traveller's Perceptions

A sample of travellers should be approached either before and/or after they have been processed by the automated border control system or by the manual Primary Line and asked whether they would be willing to answer a few questions about traveller processing. Interviewers should be positioned both at the start of the queues for each system and within the baggage hall to monitor whether travellers processed using the automated border control system successfully, attempted to use the automated system but were unable to do so, or simply chose to use the

---

[10] To provide an acceptable comparison, an equivalent manual Primary Line population to that assessed with the automated border control system would need to be obtained (e.g., similar nationalities, similar primary language etc).

manual Primary Line. Attempts should be made to obtain responses from approximately equal numbers of travellers from each of these three categories to provide a reasonable representative sample.

It should be noted, however, that this type of assessment may be problematic because travellers may be tired after a long flight and may not be willing to participate in questionnaires before collecting their baggage or may not have English as their first language. Additionally, it may be logistically difficult (costs involved, operational environment may prevent it) having a large number of individuals on-site to collect data. Previous experience, however, has suggested that travellers are generally receptive to such activities, particularly if the system is new and they found the experience of processing with an automated biometric system novel. Traveller's individual responses to all items on the questionnaire and any qualitative responses should be entered into a spreadsheet for subsequent analysis.

5 Analysing and Interpreting the Data Output

Once the data has been processed it may be analysed and then interpreted. This section describes the analysis and typical output of the data collection process[11].

5.1 Ground Truth

Ground truth is data that can lead to an understanding of the conditions where the trial was conducted. The data should be recorded over the length of the trial as it can aid in the interpretation of findings. Where the data is relatively constant, as would be expected with temperature and relative humidity indoors, just the value and the range should be recorded. Other data such as ambient illumination level should be plotted over time. Since ambient illumination levels have the potential to affect facial recognition matching, the correlation between illumination level and match scores and the number of failures may be explored. These data will be examined if, for example, it is identified that matching performance decreases, or is correlated to time of day.

Other data that should also be reported include the number of travellers per day:

- who successfully transacted using the manual Primary Line;
- who were eligible, but did not use the automated border control system;
- who attempted to use the automated border control system;
- who successfully transacted using the automated border control system;
- who were unable to complete an automated border control system transaction; and
- the above by ethnicity and gender, where appropriate.

---

[11] Please note, all data presented here is fictional and is intended for illustrative purposes only.

5.2 Matching Performance

As described in Section 3.2, the two primary matching metrics for biometric matching algorithms are FMR and FNMR. The most common way of reporting these results is via a Detection Error Trade-off (DET) plot. The DET plot is used to show FMR plotted against FNMR on a logarithmic scale. The advantage of DET over the previously used Receiver Operator Characteristic (ROC) plot is that the logarithmic plot of the DET allows better investigation of similarly performing data sets. Some fictitious data are shown in the DET plot in Figure 2, where the better performance is at the bottom left corner of the plot. The data can be displayed as overall performance, or broken down into demographic (or other) categories.

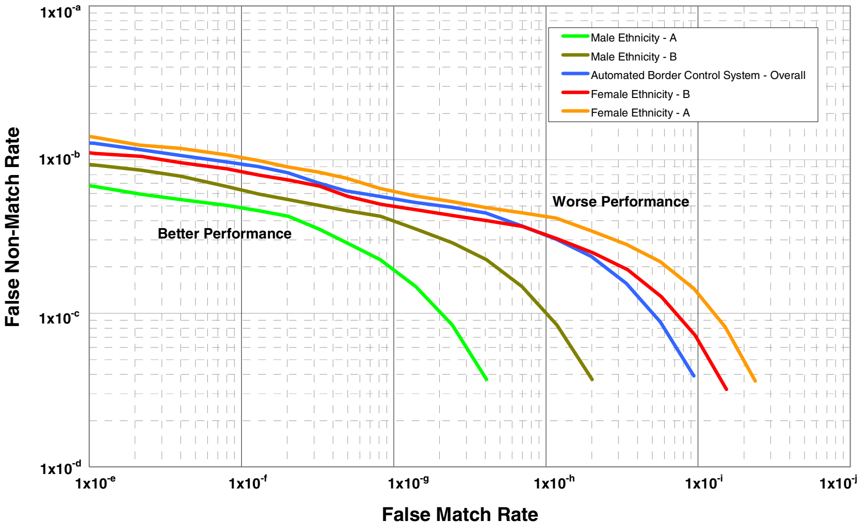


Fig. 2. Detection Error Trade-off Sample Plot

Whilst the DET shows which data set is better performing, it does not easily show the cause of the differences in performance. To better illustrate this, a cumulative probability plot that plots the match score against probability for both FMR and FNMR is also used. A fictitious example of a cumulative probability plot can be seen in Figure 3. In Figure 3, it can be seen that the Male ethnicity A is the best performing subset in terms of FMR, and the worst in terms of the FNMR. This is often the case where there is a trade-off in performance for the two metrics.

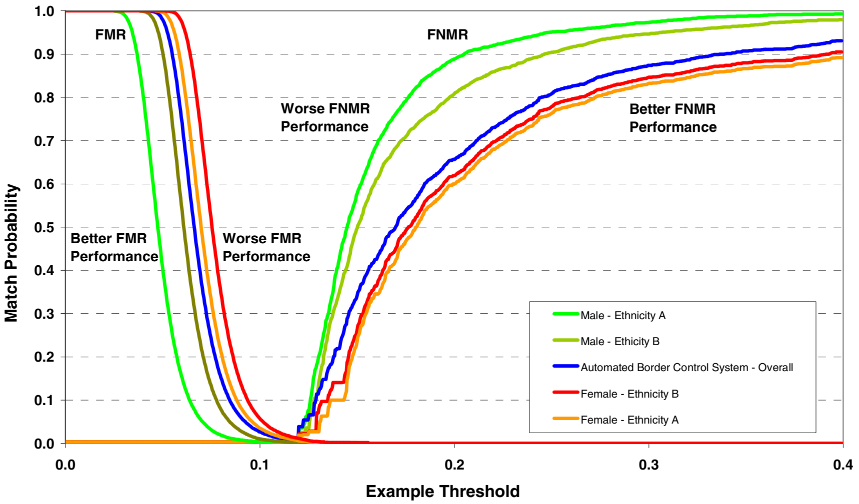


Fig. 3. Cumulative Probability Sample Plot

Matching accuracy for the human operators would likely be in the form of a table. An example of how to present this type of data is shown in Table 2.

Table 2. Sample Human Operator Matching Performance Table

Operator	*Number of Travellers*	*Correct Match*		*False Match*[1]		*False Non-Match*[2]		*Discrimination Measure (A')*	*Bias Measure (B'')*
		N	*%*	*N*	*%*	*N*	*%*		
1									
2									
3									
Overall									

[1] Where False Match equals a fraudulent traveller using a non-matching passport.

[2] Where False Non-Match equals instances where a traveller who is refer for further investigation actually has a passport with a legitimate matched photograph.

5.3 Process Timings

For both the automated border control system and the manual Primary Line processes, data should be analysed to generate overall processing times and times for each relevant sub-process (as described in Section 3.3).

The overall process timings (taken as the time for one traveller to be processed through all parts of the system from entering the processing queue until exit into the baggage hall) would allow the comparison of traveller processing times between the automated border control system and the manual Primary Line system. The evaluation of individual sub-process times may also identify if there are any processes that stand out as being excessively long which may require closer examination. For example, sub-processing times may indicate that traveller interaction with the passport reader was the longest process, which may suggest that this process should be examined in more detail to determine whether better instructions or other interventions may improve traveller interaction with this system (and thus process times).

Processing times may also be evaluated according to particular demographic or other characteristics that are thought may have an influence on processing to determine whether this is the case. This assessment may highlight that particular groups of travellers (for reasons, which should become clearer when combined with observations of the system) have difficulty or take longer to use (in particular) the automated border control system. This knowledge (when combined with observations of the system) may identify ways in which the system could be modified to better suit the needs of particular categories of travellers.

These results may be presented as descriptive statistics (i.e., minimum, maximum, median, mean and standard deviation times) for overall processing times or for relevant traveller groups. Additionally, graphs may be used to demonstrate the distribution of the sample and highlight particular trends. For example, a histogram may be used to present the distribution of processing times of the overall sample or for travellers that were and were not successful at transacting with the automated border control system. An example of this type of graph with fictitious data is shown in Figure 4.

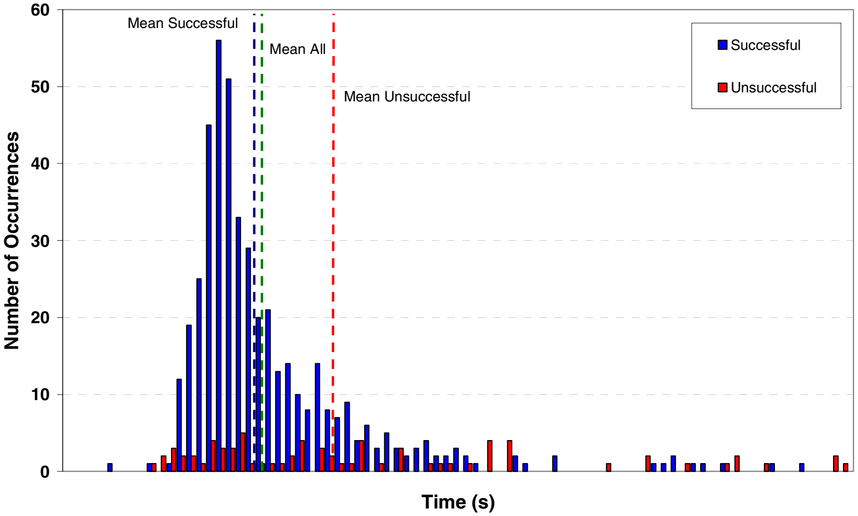


Fig. 4. Processing Time Histogram Sample Plot

A bar graph may also be used to demonstrate differences in mean (or median, if applicable) processing times between particular population groups. An example of this type of graph with fictitious data is shown in Figure 5.

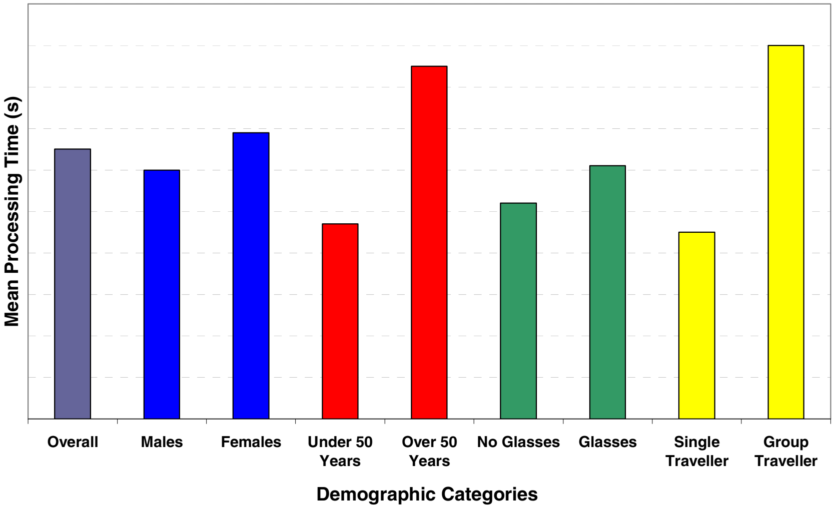


Fig. 5. Mean Processing Time for Users of an Automated Border Control System According to Relevant Demographic Categories Sample Bar Graph

A pie chart may also be used to illustrate the relative proportion of time individual sub-processes contribute to overall processing time. An example of this type of graph with fictitious data is shown in Figure 6.

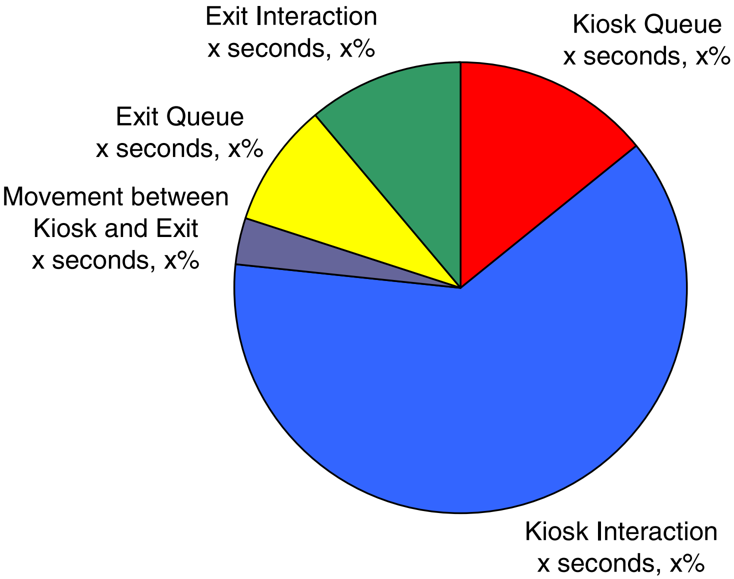


Fig. 6. Individual Sub-Processing Time Sample Pie Chart

Additionally, inferential statistics[12] may also be conducted on available processing times to determine whether there is a statistical (and practical significance) between two (or more) groups of timings. For instance, comparisons could be made between overall processing times for travellers using the manual Primary Line versus travellers using the automated border control system. This may address considerations of whether the introduction of the automated border control system conferred a processing time advantage over the previous manual system. Other comparisons that could be made may include overall processing times between particular demographic categories (e.g., travellers that processed singly versus individual processing time for travellers in a group). This may determine whether particular groups of travellers might be expected to take longer to process to inform business practices or perhaps offer modifications to the system to potentially improve processing times for these travellers. This information could, at its simplest, be used to inform resource decisions, or whether helpers are required at the automated system. Finally, this data could be used as input into a business process simulation to examine the performance of different configurations of the system under different levels of passenger movement.

5.4 Observations

For both the automated border control and the manual Primary Line systems, observational data should be obtained for each transaction for which process timings are collected. This data would be used to identify the particular (demographic or observational) subgroup from which process timings should be analysed (as was described in Section 5.3).

In addition, common mistakes travellers make or system errors would be collected in conjunction with traveller's demographic details and process timings from direct observations or from examination of system logs. This data could be used to determine the:

- proportion of travellers that were (or were not) able to process without error;
- particular groups of travellers that may have problems using the automated border control system
- frequency of each type of errors;
- the relationship between particular observational categories and:
    - processing times;
    - the failure to process using the automated border control system entirely; and;
    - the affect on biometric acquisition and quality.

This information may guide areas which require interventions to reduce the frequency of errors and improve traveller interactions with the system. For instance, it may be identified that travellers do not know where to proceed after collecting their ticket from the Kiosk or they do not understand what steps they need to take to process at the Exit.

---

[12] Inferential statistics refers to statistics, which are used to make inferences about the population from which the same was drawn [10].

The solution here (amongst other things) may be to revise the instructions and information presented to travellers to make the process clearer to them.

5.5 Traveller's Perceptions

Traveller responses from questionnaires/interviews on their use (or not) of the automated border control system should be evaluated to assess for any particular trends. Depending on the type of question and the response required, the data may be used to determine the proportion of travellers that expressed a particular view (percentage of 'yes' or 'no' responses), or for Likert scale type questions the average level of agreement with various questions across travellers. Questions that elicited a free-text response (such as on reasons for choosing to use or not use the automated border control system, or features of the system that they did or did not like) could be divided into broad categories and then assessed according to the frequency and proportion of travellers that expressed similar views. A qualitative discussion of the most common viewpoints of travellers, particularly in regards to reasons for choosing not to use the automated border control system or features of the automated system that they did not like may highlight usability issues with the system that could be modified to improve traveller functioning and/or perceptions.

6 Reporting the Results

After the analysis has been completed, the final stage in the process is to report the results so that they may be disseminated to the stakeholders. Reporting may be guided by the recommendations in [6], but should involve generating a summary of the results, with a particular emphasis on the assessment of whether the automated biometric system positively or negatively affects business processes and whether the system achieves fitness of purpose.

When using the data, the processing and the analysis techniques identified in this paper, the primary metrics that should be reported include:

1. a baseline of how long it takes to process travellers using the automated border control system and how long it takes for travellers to be processed using the manual Primary Line system;
2. whether the automated border control system expedites traveller processing (a comparison of processing times between the automated border control system and the manual Primary Line system);
3. the performance for how well each system correctly rejects fraudulent attempts at entry and how well each system correctly recognises bona-fide travellers;
4. whether the automated border control system maintains or improves border security (a comparison of accuracy between the automated border control system processing and the manual Primary Line processing); and
5. observations made as part of the operational evaluation that input into defining ways in which both systems may be improved, either in terms of reducing process timings and/or increasing accuracy or useability.

7 Conclusion

The introduction of an automated biometric system into an existing business practice requires an evaluation within the operational environment in which it is to be installed using participants from the population of likely end-users. It is recommended that a system-level assessment of performance that evaluates both the technical and human aspects of the current business practices (a baseline) and the introduced automated system should be completed. This assessment should be in terms of true-match and false-match rates, process timings and observations of the system in use and an evaluation of user perceptions. These two systems should be compared in terms of the primary metrics (i.e., accuracy of biometric matching, traveller throughput and human-system interaction difficulties) to determine the impact of introducing an automated system on existing business practices and to determine whether the automated system meets the expected fitness for purpose (i.e., in this case, whether the automated system expedites traveller processing and is at least as accurate as the manual system).

The methodology outlined in this chapter represents a tested protocol that DSTO has used in the evaluation of a number of automated biometric systems within different operational environments. It is hoped that other researchers may find this protocol useful for the evaluation of other similar systems.

References

[1] Kosmerlj, M., Fladsrud, T., Hjelmås, E., Snekkenes, E.: Face recognition issues in a border control environment. In: Zhang, D., Jain, A.K. (eds.) ICB 2005. LNCS, vol. 3832, pp. 33–39. Springer, Heidelberg (2005)

[2] Schouten, B., Jacobs, B.: Biometrics and their use in e-passports. Image and Vision Computing 27, 305–312 (2009)

[3] Kwon, T., Moon, H.: Biometric authentication for border control applications. IEEE Transactions on Knowledge and Data Engineering 20, 1091–1096 (2008)

[4] 2009 ePassport status: 60+ countries now issuing, `http://www.securitydocumentworld.com/public/news.cfm?&m1=c_10&m2=c_4&m3=e_0&m4=e_0&subItemID=1572` (accessed July 7, 2009)

[5] Goth, G.: Biometrics could streamline border crossings. IEEE Disributed Systems Online 9, article number 0804-o4002 (2008)

[6] Mansfield, A.J., Wayman, J.L.: Best practices in testing and reporting performance of biometric devices Version 2.01. National Physics laboratory NPL Report CMSC (February 14, 2002)

[7] Checkland, P.: Systems Thinking, Systems Practice, pp. 330. John Wiley & Sons, Ltd., Chichester (1999)

[8] Kaine, A.K.: The Impact of Facial Recognition Systems on Business Practices Within an Operational Setting. In: Proceedings of the 25th International Conference on Information Technology Interfaces, Cavtat, Croatia, pp. 315–320 (2003)

[9] Stanislaw, H., Todorov, N.: Calculation of signal detection theory measures. Behavior Research Methods, Instruments and Computers 31, 137–149 (1999)

[10] Urdan, T.C.: Statistics in Plain English, 2nd edn. Lawrence Erlbaum Associates, New Jersey (2005)
[11] Motulsky, H.: Intuitive Biostatistics. Oxford University Press, New York (1995)
[12] Graves, I., Johnson, R., McLindin, B.: Problems with False Accept Rate in Operational Access Control Systems. In: Proceedings of the 4th Australian Information Warfare and IT Security Conference, Adelaide, pp. 129–135 (November 2003)
[13] Calic, D., McLindin, B., MacLeod, V.: Methodology for the Comparative Assessment of Trained and Untrained Human Operators with Automated Facial Recognition Systems. Paper presented at the 10th Biometrics Institute Australia Conference, May 28-29 (2009)

Chapter 5

The Role of the Human Operator in Image-Based Airport Security Technologies

Ian Graves, Marcus Butavicius, Veneta MacLeod, Rebecca Heyer,
Kathryn Parsons, Natalie Kuester, Agata McCormac,
Philip Jacques, and Raymond Johnson

Defence Science and Technology Organisation (DSTO), Edinburgh, South Australia

Abstract. Heightened international concerns relating to security and identity management have led to an increased interest in security applications, such as face recognition and baggage and passenger screening at airports. A common feature of many of these technologies is that a human operator is presented with an image and asked to decide whether the passenger or baggage corresponds to a person or item of interest. The human operator is a critical component in the performance of the system and it is of considerable interest to not only better understand the performance of human operators on such tasks, but to also design systems with a human operator in mind. This paper discusses a number of human factors issues which will have an impact on human operator performance in the operational environment, as well as highlighting the variables which must be considered when evaluating the performance of these technologies in scenario or operational trials based on Defence Science and Technology Organisation's experience in such testing.

1 Introduction

The last decade has seen several significant security incidents at, or originating from, airports. This has led to an increased interest in and use of security screening applications, such as face recognition, baggage and passenger[1] screening. Whilst the development of these technologies began in the late twentieth century, the above mentioned incidents created not only a new wave of products, but also a significant increase in performance.

Airports are difficult environments to manage in terms of security, with large numbers of passengers, employees and members of the public moving in and out of the airport environment creating a challenge for security. Technology can clearly play a role in assisting in maintaining the security of the airport environment, with many technologies such as Closed Circuit Television (CCTV), firmly entrenched as part of the security framework.

The security screening technologies under consideration here, namely face recognition, baggage and passenger screening, are either currently used in airports or

---

[1] Although we term this type of technology "passenger screening", it should be noted that non-passengers entering the airport may also be subject to screening of this type.

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 147–181.
springerlink.com

under consideration for future introduction. Furthermore, they all rely on a human operator (or screener) to ‘interpret’ and make a final decision about the status of the person or object based on one or more still images. As such, the role of the human operator in these systems is critical.

Defence Science and Technology Organisation (DSTO) has, over the last decade, been developing methodologies for evaluating different technologies in scenario and operational testing, with an emphasis on the involvement of the human operator and the role of the operator in operating the system. Due to the nature of the technologies and the airport environments, it has been necessary for this research to take two different forms.

Firstly, research has considered the general role of humans in the operational environment through an examination of business practices amongst the security personnel of the airport [1]. Such research is generally qualitative in nature and may consist of interviews and process mapping tools to understand current procedures. This may include the identification of how responses are made to potential threats in the airport, how information is communicated between different stakeholders and how new screening and detection technologies may be integrated into current business practices.

Secondly, quantitative research on system performance and functionality has also been conducted into cases where technology forms part of a system in which the human operator also plays a crucial role [2,3]. In some cases the interaction of human and machine are inextricably bound; for example, the output of face recognition systems may be double-checked by an operator who manually checks alarms for correctness. In other systems, where the technology does not generate an alarm (or where the automated alarm system is unreliable), the final output of the system may be in the hands of the human operator who has to interpret (usually visually) the output of the technology and determine whether an alarm should be raised and what this alarm should indicate.

This chapter, based on DSTO’s experience from evaluating the performance of these systems in scenario tests and operational trials, discusses a number of human factors issues which have an impact on human operator performance and/or business practices in the operational environment. It also highlights the critical issues which must be considered when evaluating the performance of these systems prior to deployment.

2 Image Based Security Technologies in the Airport Environment

Before focusing on the role of the human operator within security screening technologies, it is important to establish the working environment of the technologies in question and also to provide brief descriptions of the technologies themselves.

2.1 The Airport Environment

In the airport environment, the baggage and screening systems are most likely to be implemented in a dedicated screening area. All individuals (i.e., outgoing passengers, airline and airport staff) wishing to enter controlled areas of an airport (i.e., any area where access to aircraft is provided) must first pass through a screening area. Personnel at each screening point must inspect all individuals and their possessions for prohibited items before allowing them to proceed.

Five screening functions are typically performed at the airport screening area: X-ray screening of property, walk through metal detector screening of individuals, hand-wand or pat down screening of individuals, physical search of property, and trace detection of explosives. In a typical situation, individuals are only required to submit their property for X-ray screening and to walk through a metal detector. However, if an alarm for a particular individual is raised at the X-ray machine, at the walk through metal detector, or they are selected for further evaluation (e.g., because of suspicious behaviour or random selection), the individual and/or their possessions may receive additional screening [4].

The basic procedure for screening individuals and their baggage is as follows:

1. Each individual places small items located on their person in a tray and then places the tray and any hand luggage items on the X-ray conveyer belt under the supervision of security staff (who will ensure that items are placed appropriately).
2. Individuals then enter the walk through metal detector under the supervision of a security staff member while their baggage passes into the X-ray machine for evaluation.
3. If the metal detector activates an alarm, individuals may be advised to remove forgotten metallic items from their person and place them on a tray for screening in the X-ray baggage system and may be subjected to a hand-wand or pat down search of their person and/or another walk through the metal detector once the offending material is removed.
4. Meanwhile an operator (possibly two) watches the display screen of the X-ray machine and decides if there are any threat items displayed on the screen.
5. The operator is in control of the conveyer belt, stopping the movement of containers for more detailed examination if required and is able to manipulate settings on the X-ray machine to enhance the presented image if necessary. Additionally, if the operator requires, any object may be resubmitted through the machine at a different orientation to aid in assessment.
6. When the operator(s) is certain that there is nothing suspicious within the item, he/she will allow the container to be moved from the conveyor belt for passenger collection.
7. However, if the operator decides that there is a suspicious object present, he/she will hand over control of the container to other security staff for follow-on action (e.g., questioning, physical search or trace detection of explosives), which will involve both the object in question and the individual who owns it.

Additionally, in some airports, the following procedure may also occur:

8. Simultaneously, while the individual's baggage and possessions are being screened, the individual will be directed to walk through a passive or active passenger screening walkway in the field of view of several cameras mounted at each end of the walkway (providing simultaneous front and back views)

under the supervision of a security staff member. Each passenger will be required to wait until the preceding passenger has cleared the walkway before they are allowed to proceed.

9. Once in the walkway a security staff member will instruct individuals to stop mid-way along the walkway and raise their arms for screening. While the individual is stationary, front and back view images will be presented to an operator (possibly two) viewing the two computer screens (front and rear).
10. The operator(s) will assess the images and, if there is no problem, indicate that the passenger can proceed through the screening zone.
11. If something suspicious is found, the individual will be asked to explain the anomaly. If this is done to the satisfaction of the security staff member, the individual will be allowed to exit the screening area. If not, the operator will hand over control of the individual to another security staff member for follow-on action.

In terms of face recognition, there may be numerous areas in an airport which could be viable for acquiring facial images (including the screening area discussed above), particularly when individuals are moving through a created choke-point, such as the walk through metal detector. The suitability of locations will depend not only on the purpose of the system (i.e., whether the objective is to detect persons of interest who are travelling, or broader attempts to consider all individuals who may be present), but also on various environmental factors.

2.2 Image Based Security Technologies

As previously mentioned, a common feature of the image based security screening technologies considered in this chapter (i.e., face recognition, baggage and passenger screening) is the role of the human operator. In fact, based on DSTO scenario tests and operational evaluations, these systems cannot currently perform at acceptable levels without a human operator to either verify or interpret the output of the technology.

In addition, all of these technologies could be classified as ‘one-to-many’. That is, the technology acquires an image (i.e., of a face, item of baggage or person) and asks the human operator to determine whether or not the image corresponds to any of the persons in a pre-defined watch list or is considered an item of contraband. In some face recognition cases not all images are provided to the human operator, with a decision making process filtering out those images which the technology deems to be ‘non-matches’ based on various thresholds/parameters. Another common feature of these technologies is that multiple presentations of the same item may not produce the same image and hence result (i.e., a person’s facial expression may have changed or a contraband item may have shifted within a bag or on the person and may therefore be more or less obscured).

In the following sections a brief description of the various technologies is provided.

2.2.1 Face Recognition Systems

Face recognition systems are currently used in airport environments to either verify that a person is the rightful holder of their identity documentation (e.g., [5]) or to identify persons of interest in the airport environment. The identity verification application is semi-automatic, and only requires human intervention if a user is rejected by the system. However, the identification of persons of interest requires human operator classification of all system alarms to resolve misclassification issues. Since the verification application in an airport environment does not require a human to assess an image acquired by the system (since, when required, the human would compare the passenger with his/her passport image) this application is not considered further in this chapter (see Chapter 4 for further information on this application).

In a face identification activity, the face recognition system processes images gathered from passengers as they pass through acquisition zones, looking for faces that 'match' with those stored in a watch list of persons of interest. Each acquired image is compared against every image in the watch list and a score is generated. Typically, if the score is above some pre-defined threshold then the comparison is considered to be a match. When a match is found, the system will signal that there has been a match and it will then display the matching images for interpretation by a human operator. This step is required since the current state of face recognition technology is such that a system alarm cannot be guaranteed to be correct. In fact, with large watch lists, it is possible that a significant percentage of individuals could generate matches and some could generate more than one match. Thus, the role of the human operator is to visually compare the acquired image(s) with a watch list image(s) and decide if the match is valid.

In this type of application, the face recognition system actually performs a sorting function and only supplies the human operator with matches that are likely (the human operator could not possibly look at all persons passing through an airport, and compare their faces with all persons in a large watch list). The human operator then performs the final decision making and initiates the follow-on action required.

As previously discussed, there is some flexibility in where face recognition identification systems can be used. These systems do not necessarily require interaction from the individuals passing through the airport, and hence can be unobtrusive or covert. It is also possible that the identification function could be incorporated as part of the border entry/exit control system. For example, when a passenger presents their passport/token and biometric sample to verify that they are a legitimate passenger, the system could use the same biometric sample to check whether or not the passenger is a person of interest. This could allow for the evaluation of the identification function as part of an evaluation of the border entry/exit control system (see Chapter 4).

2.2.2 Baggage Screening Systems

X-ray baggage screening systems are currently used at virtually all airport screening points to detect the presence of contraband (such as metallic threats, drugs or explosives) in baggage or personal belongings. The X-ray screening technologies currently available include transmission, backscatter, dual energy and computed tomography X-ray approaches.

All of these technologies involve irradiating the object to be investigated with X-rays. X-ray energy photons, when they encounter solid matter, may be absorbed, pass through the material (transmitted) or deflected (scattered). The propensity of X-rays to these three outcomes is determined by the energy of the X-ray and the bulk characteristics (density, mass absorption coefficient and effective atomic number Z_{eff}) of the material being X-rayed [6]. For instance, materials with a low Z_{eff} number (i.e., organic material) produce large amounts of backscatter radiation, whereas materials with a high atomic number (i.e., nearly all metals) absorb the majority of X-rays and therefore produce little backscattering [7].

When an object is being X-rayed, the extent to which these X-rays are transmitted or backscattered from the scanned object is collected and these are used to produce an image of the scanned item for interpretation [7]. Depending on the type of X-ray technology, this information may be presented in black and white scale, with heavy areas appearing darker than lighter organic elements [8]. Alternatively, artificial colours may be assigned to different materials on the basis of measured Z_{eff} values. For instance, low Z_{eff} number materials are typically coloured orange, while high Z_{eff} number materials are typically displayed as green [7]. Additionally, properties of the scanned material, such as calculated mass density, nitrogen content and Z_{eff} number, may be generated for each type of material to aid in identification [6].

It is argued that the characteristics of explosives, while not unique, are sufficiently different from everyday objects that the detection of similar properties is suggestive of a high probability of the presence of explosives in a particular item [6]. It is for these reasons that some X-ray screening systems are able to implement the automatic detection of explosives (and explosive-like materials). However, regardless of the level of automatic detection all systems still rely, to some degree, on operator recognition of dangerous objects/materials by way of evaluation of shape, colour (density) or texture of objects in the image to make the final decision.

2.2.3 Passenger Screening Systems

Passenger screening technologies must be capable of detecting objects or materials of interest that are hidden underneath clothing or otherwise concealed on the bodies of persons. A wide variety of technologies are employed, although the most common technologies involving a human operator include backscatter, passive mm-wave, terahertz and active mm-wave systems.

2.2.3.1 Backscatter X-ray Systems

Backscatter X-ray systems involve the use of X-rays that penetrate through clothing and other low density material and (at intermediate energies) are scattered back efficiently from low Z_{eff} number materials under the clothing. Most explosives are good backscattering agents, so they show up as being brighter than human skin. Other materials, such as metallic weapons (i.e., guns, knives) are not efficient backscattering agents and show up as darker than human skin. In either case, hidden materials and objects may be identified by way of a contrast image and the nature of the explosives may be inferred from the positive contrast they produce.

In practice, individuals will stand in front of an imager and be scanned by a moving beam/detector combination. X-rays are collimated and scanned using mechanisms made from lead or similar high attenuation materials. The scanning process takes time and the individual must be scanned in both front and back views, so throughput is relatively slow. This type of system would only be considered for screening individuals who were suspected of concealing an item of interest and most individuals would not be screened in this way.

Imagers of this type are capable of producing high resolution images of the unclad body, so privacy may be a concern and privacy filters may be required to protect individual's modesty. These machines also irradiate individuals with very small amounts of ionising radiation, so health issues may need to be considered.

These machines have some automatic detection functions, providing a threat indication based on the effective atomic number of the detected materials. However, the primary classification is provided by a human operator who views the images from the system and decides on the position and nature of any hidden objects.

2.2.3.2 Passive mm-wave or Terahertz Systems

Passive imaging systems, such as mm-wave or Terahertz (THz), are designed to produce body images from the thermal radiation passing through clothing. Any object concealed under the clothing will produce a contrasting image if it is at a different effective temperature to the surrounding body. In use, the imagers could either be set up to scan individuals as they walk through a screening zone. Individuals may be asked to stop in an appropriate pose as they are scanned, or (unlike the other passenger screening technologies) could be scanned unobtrusively or covertly at some other location. In either case, the throughput can be relatively high and this type of system would be expected to be used to screen all individuals.

The image quality from passive imagers is relatively poor since only a small amount of thermal energy is available at these frequencies. The resulting images do not include much detail and may be noisy, so a human operator is required to detect and identify threats, based on appearance and placement.

2.2.3.3 Active mm-wave Systems

Active mm-wave scanning portals generate images of the bodies of individuals through their clothing. The sector scanning mechanism allows a full 3D image to be generated and this can subsequently be viewed from any angle in order to assess the nature of objects detected.

In use, individuals are asked to enter the portal, stop in a required pose while they are imaged and then leave the portal. An operator then assesses the imagery produced. This process can be relatively quick if required, so this system could potentially be used either for screening all individuals or alternatively for screening only those individuals who are selected as being suspicious by other means.

This type of system does not provide any assistance in identifying objects, so the sole classification mechanism is provided by the human operator, who must detect and identify threat by way of their appearance and position in the imagery.

3 The Role of the Human Operator within the System

The human operator is defined as the individual within the airport environment working with the technology who makes a decision regarding identification of a threat based on an image of the potential threat (i.e., face, item of baggage or person). If necessary, it is the human operator's role to escalate the matter within the system so that the relevant person(s) within the system can take an appropriate course of action. Human operators may also be required to operate and monitor the technology and undertake other tasks not related to the operation of the technology.

Although, as stated above, the human operator represents only one element of the system, the human operator can be considered the key element in the system as they make the ultimate decision about whether a person of interest has been identified or a contraband item has been found within baggage or on the person. Based on DSTO evaluations of this type of technology, without a capable human operator, current image-based airport security technologies cannot achieve levels of performance which are useful in the operational environment.

It is for this reason that the technical and human elements of airport security systems must work together effectively, as weaknesses in either element diminish the effectiveness of the overall system. Such limitations need to be understood and accounted for in the development of policy and procedures governing the operation of these systems. However, while much headway has been made towards understanding and improving the technical elements of airport security screening systems (e.g., the algorithms, imaging techniques and database construction), comparatively little attention has been paid to understanding the role of the human operator [9, 10].

3.1 Automation of Security Tasks

One common feature of the security screening technologies discussed here is that they all automatically acquire the image of the potential threat to assist the human operator.

Systems are often automated to increase efficiency and effectiveness, reliability, and safety (see Section 3.1.1). A number of levels of automation have been defined, ranging from complete manual control, whereby the automation offers no assistance and the human is in complete control, to a system completely automated where the human is not informed about decisions made by the system and is largely ignored [11].

It is argued, however, that when included in the process, human operators add judgement to the system and are better able to respond to unforeseen changes than technology alone [12]. Security screening tasks require high levels of cognitive ability and a human operator is often limited by the capacity of their working memory [13]. In this way, security systems with a need for high volume cannot perform effectively without some automation. Nevertheless, care must be taken to closely assess the role of the human in the system in order to reduce the likelihood of error. Errors associated with system performance can originate directly from system functioning or may be the result of human issues linked to attention, perception and cognition in operating the system. Although system failures and breakdowns cannot be stopped, human errors can be managed and addressed to further develop connectivity between the human and the machine in the system and to stop further malfunctions from occurring [13].

3.1.1 Reasons for Automation

There are a number of different reasons for automation [13]. A process may be automated because it is impossible or hazardous for a human to perform. For example, it is not possible for humans to replicate the X-ray task, nor is it possible for humans to replicate the task of matching individuals against thousands of facial images. Also, given the nature of some of the potential threats (i.e., explosives, weapons), it could be hazardous for humans to undertake the equivalent task.

Automation may also be employed for difficult or unpleasant tasks ordinarily carried out by humans that may be very challenging to undertake and/or may lead to lower rates of performance. Such tasks often require vigilant monitoring for very rare events and consistent decisions to be made, both of which are difficult tasks for a human alone [14] (i.e., physically searching passengers).

Additionally, automation may provide a system which simply extends the human capability to do the same thing in complex scenarios. Automation in these circumstances relieves the operator from cognitively demanding tasks and excessive working memory requirements, while also facilitating a multitasking capability. For example, in the case of face recognition screening applications, research suggests that humans are highly skilled at distinguishing faces (particularly involving familiar faces [15]), but that they are poor performers in face recognition tasks involving unfamiliar faces [16]. Further, reviews of the literature have shown that factors such as race, disguises, changes in appearance and quality/type of image can impair human recognition accuracy [17]. Overall, the face recognition literature suggests a variable rate of human performance for familiar faces [15] and very poor performance for unfamiliar faces [18]. In this case, automation provides an extension of the human ability to recognise faces, which is ordinarily a cognitively demanding task when the operator is provided with many unfamiliar faces and asked to quickly make a decision. In addition, detection systems that automatically detect threats and highlight them for an operator's consideration extend human capability in interpreting output from baggage and passenger screening systems by highlighting possible threats for more in-depth consideration.

3.1.2 Limitations of Fully Automated Systems

In addition to the need for a human operator to manually check identified alarms or review X-ray images of baggage or people for contraband in airport settings, there is also a need to observe for suspicious behaviour. The screening technologies examined in this paper do not currently undertake automated evaluation of human behaviour, although such systems are currently in development [19].

Humans are able to naturally detect or be trained to attend to the voices, words and actions of individuals indicating dishonest behaviour. The face provides valuable information regarding others' emotional states and most individuals can accurately interpret these. More specifically, individuals can accurately link increased facial movement or blinking to stress, anxiety or nervousness [20], and can use body and voice cues to distinguish defensive behaviour or fear [21]. Additionally, humans can utilise their abilities to carry out important confirmatory tasks, such as interrogating an individual, establishing travel history, or checking the validity of identification documents.

A screening system which combines the strengths of technology and automation with human perception is likely to be best in optimising performance, rather than a solution based solely on one or the other. Ideally, an overall system which takes into account the limitations of both the human and automated components of the process, in order to provide optimal performance within the limitations of the scenario, should be deployed.

3.2 The Impact of the Human Operator on Overall Performance

It is clear that human operators contribute to the successful performance of security screening systems. However, understanding the way in which these security screening systems operate is crucial to understanding the improvement in performance a human operator can provide.

3.2.1 Consequences of Non-optimal Decision Thresholds

In airport security settings, correct identifications are crucial. For instance, failure to detect a person of interest or something hazardous may result in a security breach. Conversely, unnecessarily inconveniencing passengers negatively impacts on business practices and may lead to unwanted media interest and public concern.

Consequently, the decision thresholds need to be set at a practical and operationally viable level. Taking into consideration the trade-off between the number of incorrect alarms and missed threats, setting the system's decision threshold too high can lead to correct identifications being missed. However, setting the threshold too low can lead to a high number of false alarms, causing possible overload of the system, particularly the human operator. Allowing a high number of false alarms could increase workload for airport personnel (e.g., unnecessary questioning of innocent persons). A high number of missed detections will increase the risk that threats will not be detected, which may have security implications. Therefore, a completely automated system would require a threshold setting that would compromise either the number of incorrect alarms or missed threats, often to the point of making it operationally unviable.

The addition of a human operator in the system enables these threshold settings to be tailored to achieve the highest level of performance. That is, the threshold setting for an operator based system can be chosen to allow for an appropriate number of false alarms which the operator is able to deal with quickly and effectively, while generating a higher number of correct identifications and reducing the number of missed detections.

3.2.2 Limitations to the Impact of the Human Operator

Automated systems affect the performance of the human operator in four key ways.

Firstly, if the operator is not shown a threat they are unable to make a positive identification. For example, as is the case with some face recognition systems, only watch list images which achieve a score above a certain threshold setting will be shown to the operator. Therefore, it is possible that if the matching watch list image does not score highly enough this image will not be presented to the operator as a possible match. Additionally, for all systems, if the screening equipment does not provide an image of the threat (e.g., if the threat is obstructed by the presence of other objects or due to poor placement) the operator is unable to make a correct identification. In essence, if a threat is not shown, it cannot be identified.

Secondly, if the operator is shown an image of inadequate quality to make a correct decision regarding the likelihood of a possible threat this will compromise the ability of the operator to perform at an adequate level. For example, research into human face recognition suggests that humans are vulnerable to the effects of viewpoint pose and illumination, so much so that if the two images under examination are in different poses, or if the two images are taken under vastly different lighting conditions, the human operator will have significant problems identifying them as matching [22]. Similarly, humans have difficulty reaching conclusions about unclear X-ray images [23].

Thirdly, image-based airport security screening technologies rely on a certain level of operational proficiency and may be affected by system failures or environmental issues, such as breakdowns from dealing with large amounts of input (e.g., images, identifications) or possible power outages. Essentially, the operator has no real control over system functioning or breakdown and in these cases cannot reach a decision if the system is not operational. Indeed, one of the functions of the operator should be to monitor the system connectivity and functionality, and initiate repairs to the system if required.

Finally, the system needs to be appropriately calibrated to allow for the human operator to operate at an optimal level of performance. This calibration should recognise that the capabilities of different human operators will vary. Although the human can contend with a certain number of false alarms, overloading the operator can lead to poor decision making and failures of the system to extend the capability of the human. Humans will be fallible at some point, but the system should work with, rather than against, the operator to provide the highest level of performance.

4 Factors That Impact Human Operator Performance

In this section the focus is on the impacts on human operator performance as discrete from those associated directly with the system functioning.

4.1 Usability

Human operator performance can be impacted by the usability of the security screening system that they are required to operate [24]. A system is considered usable if the intended operators can meet a desired level of performance operating it, the amount of learning or practice to achieve this level of performance is appropriate, the system does not place any undue physical or mental strain on the user, and users are satisfied with the experience of interacting with the system [25]. In assessing the usability of a system the focus should, therefore, be on the operator's appraisal of the system and include an objective assessment of the effectiveness and efficiency of the system, as well as the operator's overall level of satisfaction with the device.

A number of factors impact on the usability of a system, ranging from the ergonomic design of the work area, seating and the work station itself; to the efficient design of the input (e.g., control consoles) and output (e.g., display screen) devices, as well as user manuals. Input devices on security screening systems can vary in design, from a standard keyboard to a proprietary designed console specific to that screening system. The output device should be large enough to aid the operator in their decision making [26]. If a large amount of information is required to be displayed, the use of additional screens should be considered as presenting images as large as possible will aid the operator in analysing the image and reaching a decision. Graphical User Interfaces (GUI) that use familiar designs (icons and actions), are readable (in terms of text size and colour), and which utilise logical information placement and information flows can enhance human operator performance by reducing stress and fatigue [27].

It is important that usability issues are considered early in the system design process to ensure that the impact on human operator performance is minimised, and unnecessary costs (e.g., to modify equipment post deployment) are avoided [28]. Unfortunately many such systems are designed without consideration for human factors issues, which often force operators to use poorly designed systems that directly impact on performance.

4.2 Training

Training is considered to be a key area that impacts on operator performance, however, the link between training and performance is poorly understood. Training given to operators tends to vary, with some operators given computer based training, while others only receive on-the-job training.

In terms of human face recognition performance there remains disagreement regarding the impact of training in recognising unfamiliar faces. Experience and training are both central to the performance of human operators [15], however, results on the usefulness of training are inconsistent. Several studies conducted in the past have identified that training does work [29, 30], whilst others have found that training has no effect on participants' subsequent performance [31]. There is difficulty in identifying the most effective method, scope and duration of training to produce the best results [15].

The results of DSTO trials of baggage and passenger screening systems suggests that differences in training packages may influence or bias the operators' performance [32]. For instance, if training packages tend to only show images of individuals or baggage with contraband, the high presentation of threat items could lead to an unconscious response bias towards raising an alarm. Similarly, if users are only trained on images of individuals or baggage with contraband, their discrimination (i.e., ability to identify a non-carrying individual), may be lower. Hence, when providing training for airport security screening technologies, it is important to train human operators with examples where a threat item is absent, as well as examples when a threat is present.

The detection of contraband items through X-ray machines is quite complex and challenging as there are multitudes of different contraband items that come through X-ray devices in various configurations and viewpoints. Additionally, a change in orientation can make an object harder to detect. In reality contraband items are very difficult to recognise without any training [33].

Feedback is used in many training situations and it is a process by which information about performance is provided to an individual or group in order to facilitate learning, with the overall aim of improving future performance [34]. Feedback can prove to be an effective training mechanism as demonstrated in Chapter 6. However, although it is possible to use immediate feedback when training operators on X-ray technologies, this is not always the case with all other forms of screening devices in airports.

The majority of previous research has been limited to the detection of contraband through the use of X-ray technologies [35, 36, 37], so further research is required to measure the impact of training on operator performance with each new type of screening technology.

4.3 Workload

Since the human operator plays an important role in the systems considered in this chapter it is crucial that the relationship between their workload and performance is not only understood, but managed. It is also important to recognise that the abilities of individual human operators will vary. For example, some operators may be better able to cope with a high volume of alarms, but may be less capable of multitasking.

4.3.1 Number of Alarms

In systems where the human operator provides the final decision making function, the operator work-rate will depend on the number of alarms. Since the apprehension of persons of interests or those carrying real threats is a rare event, the primary source of alarms will be false alarms. In the case of a face recognition system with a large watch list, a significant percentage of individuals may generate false alarms, with some individuals possibly generating matches with multiple watch list identities. The workload on the human operator may therefore need to be managed by restricting the passenger flow rates.

In the case of baggage and passenger screening systems, where the human operator provides the classification of objects imaged through bags or clothing, the false alarm rates from innocently carried items and clothing could also be a significant fraction of the passenger throughput rate.

Considering the vast majority of alarms will be false and there may be many such false alarms, operators and observers will need to be trained to have the expectation that, although most alarms will be false, they still need to remain vigilant in order to detect the rare true alarm. Since the natural human tendency is to ignore the source of continual false alarms, training to counter balance this tendency will be essential.

4.3.2 Multi-tasking

In some circumstances the screening function may not be the only responsibility of the human operator. Most research in the area of multi-tasking suggests that the introduction of subsequent tasks distracts from the original task, impacting negatively on performance [39].

As an example, for face recognition systems the human operator may be located in a CCTV control room, monitoring the CCTV feeds generally until a face recognition system alarm occurs, or the throughput through the capture zones increases. For baggage and passenger screening systems, a secondary responsibility of the human operator may be to monitor any unusual behaviour in the area.

Where human operators have multiple roles there are several potential implications. Firstly, the multiple roles may have an impact on the time taken to review the alarm/image by introducing a ‘time delay’ not related to the review process itself. For face recognition this could result in a failure of the appropriate response (i.e., question and/or detain the individual) to be actioned, whereas for baggage and passenger screening this would impact on the flow through rate at the screening area. A delay in reviewing the alarm/image may also cause the human operator to rush their decision which could result in an incorrect decision.

Secondly, an incorrect decision may also result if the human operator is distracted (i.e., by monitoring suspicious behaviour) whilst making their decision. For face recognition systems the distraction of other tasks may result in alarms not being reviewed in their original order, which may impact on the ability to action a response. For baggage or passenger screening systems this may result in confusion as to which bag or person contained a possible threat. Additionally, for face recognition systems, it may not be possible to review some alarms if the alarm queue is of insufficient length and/or alarms have been automatically removed from the system after a period of time.

Therefore, where human operators have multiple duties it is important that these are considered in the assessment of the systems in order to highlight any implications for the performance of the systems and also existing business practices.

4.3.3 Individual versus Team Dynamics

It is also necessary to note that there may be performance differences between individuals and teams. These differences may have important implications within the airport security domain. It could be considered that a system that requires fewer resources (i.e., one rather than two operators) would be desirable. However, given the importance of precise decision making, if an extra operator greatly increases the accuracy of the system, this could be worthwhile.

Unfortunately, there is a great deal of disagreement within research in the area of group decision making. In some instances, teams have been found to be better able to reach decisions that involve more risk than do individuals [39], whereas other teams have been found to reach decisions that are more conservative in nature [40]. A number of factors including communication skills, and shared expectations may affect group performance [40, 41]. Similarly, phenomena known as groupthink and group polarisation can also occur in team situations [41, 42]. The study of group effects is made even more complex because such effects are influenced by the personality and cognitive attributes of the individuals involved. In Chapter 6, a more detailed exposition of group effects, in the context of these individual differences, is presented in the theoretical framework of a human factors experiment of an airport security imaging device.

It should also be recognised that multiple human operators could be utilised with independent decision making functions. For example, given that X-ray images display a reasonable level of complexity and that each image typically contains multiple types of stimuli (e.g., innocent items, as well as explosives, guns and knives), which are usually presented within the image with different colours and at different orientations, two operators may be more efficient than one at interpreting these images. In a series of studies [43] it was observed that the time taken to search for two visual targets simultaneously was higher than for two single target searches, with there also being an associated loss of accuracy. Therefore, X-ray searches may be more efficient with two operators, with one searching for one type of stimuli and the other for another type of stimuli in all presented images.

Additionally, if several baggage or passenger screening systems are deployed a different human operator would be responsible for each one. Also, for face recognition systems the alarms could be split (especially in periods of high throughput) and passed onto different human operators for assessment. In all of these cases the human operators are considered as individuals rather than as a working team.

4.4 Fatigue and Stress

In [44] the authors found that a person's ability to maintain vigilance and attention reduces over time. Inattention can be induced by fatigue, which has been widely studied as the cause of errors, inefficiency and risky decision making [45, 46].

Shift workers are particularly susceptible to fatigue, mainly due to the impact of shifts on circadian rhythm. Circadian rhythm refers to the cycle of biological processes that take part in all humans throughout the day. Circadian rhythm peaks during the hours of 0800 and 1400 and considerably lowers in the evening and early morning. Personnel working within an airport security environment are usually required to work to unusual shift timetables based on the schedule for arrival and departure flights. These schedules can include working very early in the morning, working longer hours (i.e., 10 hour shifts) to allow for more consecutive days of rest between work periods, and working a mixture of different shift times to maintain a continuous rotation of schedules (e.g., early in the morning and late at night).

Although strategies are usually applied to allow workers to adapt, full adaptation is never really achieved as the individual will still be exposed to some evidence of the earth's natural day and night cycle affecting synchronisation to a new circadian rhythm. These conditions create an environment where these individuals are working when they are fatigued and when circadian rhythms are at their lowest levels, which has been shown to increase error rates [13, 47].

In fact, fatigue, stress and distraction have long been reported as key factors impacting on human operator performance across a range of settings [48]. Human operators will often be required to perform under time pressures, where there may be crowds of people. In [49] it was found that humans often perform worse when placed under a time pressure as opposed to normal conditions. Such pressure can cause stress and promote risky or biased decision making.

Additional tasks can also cause operator stress (see Section 4.3.2). The requirement to consult a decision aid or check multiple screens for information can direct attention away from the core task. In addition, the higher the state of the alert, the greater the number of false alarms presented for the human operator's consideration [50], which can subsequently impact on performance (see Section 4.3.1). An operator performing face recognition who is faced with two possible matches to decide between has to expend less effort to arrive at a decision than one which has 20 or more possible matches to decide among [51].

In addition to time pressures and additional tasks, the tasks human operators are required to complete require attention to detail and the maintenance of such attention over significant time periods. Periods of inattention can lead to significant decreases in performance where signals can be missed [44].

It is also necessary to note the importance of individual differences. That is, different human operators working in the same environment may have different reactions to stress, and/or be less susceptible to fatigue and distraction [47]. In investigating these differences further it may be possible to derive other possible strategies which can be employed to overcome fatigue associated with working in an airport security environment.

4.5 Human Operator Individual Differences

There is evidence to suggest that certain innate abilities or personality characteristics may be associated with better performance on human operator tasks [52, 53]. Knowledge of innate abilities or other characteristics important for operator performance may be used to guide personnel selection and/or training for security screening functions.

Perceptual and cognitive abilities play a role in the ability of an operator to interact with image-based security devices. However, in the context of security based devices, little research has gone into:

1. identifying what these specific abilities are,
2. predicting performance in the real-world application on the basis of these abilities, and
3. the extent to which any such abilities may be learned or innate.

There is evidence that performance on a test known as the X-Ray Object Recognition Test (X-Ray ORT) can predict performance on the real world task of identifying contraband in X-ray images of baggage [53]. This test acknowledges the influence of viewpoint of the contraband, the inclusion of multiple non-contraband items in the same bag and the effect of overlaying items from the screener's perspective on overall performance on the real-world task. It is the ability of the screener to cope with these aspects of the task that appear to predict real-world performance. It should be noted that in this test the images are very similar to what the screener would view in the real world so it gives a good practical test of performance on the task. In addition, there is evidence that an abstract test of perceptual processing ability can also partially predict performance on an image-based airport security device (see Chapter 6).

In the context of face recognition systems, there is evidence to suggest that recognising faces is innate, and qualitatively different to the recognition of other objects [54, 55, 56, 57]. There are a number of arguments in support of this assumption. Firstly, it has been proven that faces are easier to remember when presented in an upright rather than inverted orientation [57]. Secondly, there is a form of agnosia known as prosopagnosia (or face blindness) in which patients are unable to recognise faces, but yet often have no difficulty with the recognition of other objects [57]. This suggests that there is a dedicated part of the brain responsible for face recognition. Thirdly, studies have demonstrated that newborn infants can discriminate between individual faces, suggesting that face recognition is well developed by birth, and may, in fact, be an innate ability [58, 59].

In addition, to cognitive and perceptual processing ability, it has been suggested in Chapter 6 that cognitive processing styles may impact performance in screening or operator tasks. In particular, the bias towards rational or experiential processing styles have been hypothesised to impact performance in terms of whether an individual is biased more towards responding quickly or more accurately. The experiment in Chapter 6 includes the Rational Experiential Inventory [60] and that chapter describes the underlying theory in more detail.

4.6 Biases

Human operators may be subjected to a number of different biases that can directly or indirectly affect their decision making and performance.

Whilst it is important to determine if biases in the responses of human operators are present, it is equally important to understand why they exist and how they can be removed. In some cases it may also be desirable, as part of business processes, for some biases to remain. For example, it may be appropriate that human operators effectively use a lower decision threshold for items which may be explosives.

4.6.1 Ethnicity, Gender and Age Biases in Face Recognition

The ‘own race effect’, where people have been found to more easily recognise and distinguish faces of their own race more accurately than faces from other races, has been well documented [61, 62]. Nevertheless, research focusing on the own-race effect has shown that people can be trained to better recognise faces of different races through experience, exposure and practice. This was done through directing people toward certain facial cues and providing feedback on performance improvement in face recognition with own race figures [63].

Similarly, a gender bias has been observed where people have been found to recognise faces of their own gender more accurately than faces of the opposite gender [64, 65]. However, research related to own-gender bias is not conclusive, with some studies demonstrating that females are better at remembering faces and are able to outperform males in recognising male faces [66, 67, 68]. It has been argued that this occurs because females have a higher interest in people and as a result remember more faces [67, 68, 69]. While most findings do not report male own-gender bias, there have been a number which have found some effect [65, 67, 70, 71] and other studies which have reported that males recognise male and female faces equally well [67, 69, 71]. In terms of which faces are generally better recognised, mixed findings have been reported. Some researchers have found that female faces are better remembered and recognised because they are found to be more distinct [68, 72, 73], while others have indicated that male faces are better recognised than female [22, 70].

More recently, research has turned to the own-age effect. Here, researchers have found an own-age bias for children, adolescents, and adults (young and elderly). That is when making face recognition decisions, people tend to perform better when asked to make decisions about faces of their own age, than faces from other age groups [74, 75].

4.6.2 Cognitive Bias

In relation to all types of security systems, biases may relate to cognitive constraints, such as perception, judgement and decision making. These issues have been reviewed in [76]. Additionally, the role of confirmation bias [76], that is, a tendency to confirm any initial theory or preconception whilst avoiding disconfirming information in the forensic identification setting has also been studied extensively [77, 78, 79]. Research has found that decision makers have a threshold that must be reached before a certain decision, such as to identify a person or threat, can be made. Many factors can influence this, including time pressures, accountability, expectations and emotions surrounding the task [80]. In addition, it has been found that human operators often base their decision making on past experiences more so than on logic or rationality [76].

In terms of baggage and passenger screening systems, evidence suggests that human operators can only recognise something they have seen before, or is similar to something that they have seen before. Thus, operators may be able to easily recognise the presence of a gun if the image produced shows the gun in profile, as this matches the expected form of a gun. However, an image of a gun from a different angle, which does not clearly display the typical gun shape may be less identifiable to operators.

Human operators can also become biased towards raising an alarm when conditions suggest an alarm is more likely. For example, if intelligence has been received that a person of interest may be in the area, or if there is heightened awareness because a threat item has recently been found, an operator may err on the side of caution (i.e., making too many false alarms) in the hope of identifying the suspect. In other words, the amount of visual evidence they require in order to decide to call an alarm is reduced (i.e., the operator's bias may reduce). However, this change in bias does not affect the operator's ability to discriminate between contraband and non-contraband cases.

These issues highlight the importance of ensuring that staff are not only adequately trained in their role of assessing alarms/items, but that they also have a clear understanding of the business processes in place, as well as an understanding of procedures and best practices to reduce common biases [76].

4.6.3 Reliance on Automation

A number of the current security screening technologies are configured to provide varying levels of automatic threat detection. For face recognition systems this may be in terms of the presentation of match scores (or percentage of similarity) in conjunction with images for the closest matches generated by the system. In relation to baggage and passenger security screening systems this may be the highlighting of suspected explosives or other threat objects within an X-ray image on the basis of the Z_{eff} number of the materials. This level of threat indication may be a useful guide for operators when making decisions. However, it is possible that certain operators, or all operators in certain situations (i.e., high pressure, time constraint or fatigue conditions), may come to rely on these automated prompts to generate their decisions entirely.

For instance, in face recognition systems, operators may assume that if a decision is difficult or if there are time pressures that the image pair with the highest match score presented must be correct. Consequently, operators may not adequately consider another image that was in fact a correct match, but which obtained a lower match score and was therefore presented lower down in the list of possible matches for their consideration.

Similarly, as was noted in a previous evaluation of baggage and passenger screening systems [32], operators may fail to detect a threat object if it is not highlighted by the system on the basis that if the machine did not detect a threat then one must not be present. Additionally, in this same evaluation, interviews with operators after the trial indicated that some operators doubted their own judgements about the presence of a threat after the machine displayed a threat detection.

Operators indicated that this often changed their previous negative decision about the threat status of an X-ray image to a positive one.

Thus, over reliance of operators on the technology to automatically flag a threat (or not) for them to evaluate may lead to a failure of operators to adequately interpret images themselves, potentially leading to biases in decision making.

4.6.4 Consequence of Actions

The propensity of the human operator to make a particular decision may be dependent on the consequence of the action. Here the consequence of the action can include the consequence of not identifying a threat (i.e., the threat that the item poses, both in general and also to the human operator) and also the consequence of incorrectly classifying an item as a threat (i.e., the response to the action). For example, the tendency to classify an item as a threat may differ for images suggesting the presence of a relatively innocent table knife compared to a more harmful plastic explosive. In this case the threat posed by each item is different, and the appropriate response would most likely be different as well. Additionally, if the implication of a decision was to completely shut down an international airport for a number of hours, operators may be less likely to identify a threat.

4.6.5 Speed of Processing

The time required to perform cognitive tasks depends on many factors, not the least of which is the operator's requirement for speed versus accuracy [81]. The consequences of making an inaccurate (and potentially dangerous decision) must be balanced against the need to process individuals as quickly as possible to meet logistical (or administratively imposed) requirements. This can lead to a speed/accuracy trade-off. Responses made under speed stress can be faster than those observed under accuracy stress; the reduction in reaction time may be accompanied by an increase in error rates, with speed being traded for accuracy [81].

Additionally, there is also a requirement, after a decision has been made, for personnel to have sufficient time to act upon that decision. As was previously discussed (see Section 2.1) there may be numerous suitable locations to acquire images for the purpose of face recognition. In many of these locations it may not be possible to contain the individuals in the area whilst the face recognition process is undertaken. Hence the speed of the system, including the technology and human operator, is critical. For example, if images are acquired whilst the individuals are at the inwards entry point, it is important that the face recognition process is completed in sufficient time to allow the person of interest to be located before they have left the area.

Since baggage and passenger screening technologies are deployed at specific screening points (see Section 2.1) a degree of control can be achieved over individuals attempting to pass through the area. This control ensures that any threat identified by the human operator can be contained in the screening area so that an appropriate response can be initiated.

5 Recommended Methodologies for Evaluation of Systems

It is argued that, in order to evaluate security screening technologies for the airport environment that use human operators, an understanding of the complete system, including the role of the human operator is required. The evaluation methodologies available in the literature typically focus on the technical performance of the technologies and fail to account for the influence of the human operator on operational performance [see, for example, [82]).

This section provides an overview of the recommended methodologies developed by DSTO. The focus here is on the role of the human operator, with further details for evaluating specific systems contained elsewhere in this book.

5.1 Types of Evaluations

The independent evaluation of security screening technologies prior to their implementation in the airport environment is critical to ensure that an appropriate level of performance in the operational environment is achieved. In most cases a multi-phase evaluation will be required, with the role of the human operator an important consideration in all the phases.

5.1.1 Technical Assessments

Once a number of potential solutions have been identified it may be advantageous to conduct technical assessments with the various technologies to baseline the performance levels achievable.

The aim of technical assessments of face recognition systems is typically to compare competing algorithms using a common sensor in a controlled environment [82, 83]. For baggage and passenger screening technologies technical assessments in a laboratory environment may provide the only opportunity to test the technologies against the full range of items of interest. As part of this testing it may be possible to provide guidance on procedures for an operational environment (e.g., suggesting how to best place items for baggage operators).

Whilst the focus of technical assessments is the technical capabilities of the technologies, it is important that human factors issues are also considered. For example, it may be possible to eliminate from consideration technologies which are comparatively complex or time consuming to interpret, and therefore unworkable in an operational environment, or systems that clearly do not enable a human operator to undertake their role effectively.

5.1.2 Scenario Tests

The purpose of the scenario testing is to determine the system performance in an environment simulating (as much as possible) the operational environment and processes [82, 83].

For scenario tests sufficient subjects or items will be required to achieve confidence in the results, noting that, where relevant, multiple iterations with the same subject or item can be conducted. In general the role of the human operator should be an important consideration in scenario tests, with both the performance of the technology and the impact of the human operator on overall performance assessed.

Scenario testing is an important phase in most evaluations, especially if no technical assessments were undertaken. However, it is important to recognise that results will not be representative of those achievable in the operational environments due to differences in the environments (i.e., lighting for face recognition), subject behaviour and the various impacts on the human operator previously discussed (i.e., fatigue, workload, consequence of actions and bias).

5.1.3 Operational Trials

The goal of operational trials is to assess the feasibility of the complete system in the operational environment and with a target population [82, 83]. Here the system includes not only the technology but also the role of the human operator, the ability of the business processes to respond to threats and the capacity of the system to function for extended periods without system failures.

However, whilst operational trials can provide the most realistic assessment of the suitability of a system for deployment, there are typically many challenges in conducting such trials (e.g., obtaining permission to conduct an evaluation in an airport environment) (see Section 6). As such, scenario testing (in the operational environment) can be incorporated into the operational trial to mitigate some of the limitations of operational trials (i.e., sample size).

Often the ultimate goal of assessing a system is to determine the feasibility of implementing a system in a number of different, but similar, airports. As part of the assessment it is important to recognise that the different environments will result in different performance simply because of the different operational conditions and circumstances. If there is flexibility in choosing a trial location it may be advantageous to choose the most challenging environment, which may, for example, be the airport with the greatest throughput and hence pressure on the human operator.

Following the operational trials, it may be possible to reuse some of the data collected to conduct additional analyses. For example, biometric face images could be reprocessed by the algorithm against different watch lists, or baggage or person screening images could be replayed to a larger pool of human operators for evaluation.

5.2 Primary Measures of Performance

Performance measures are a critical component of any evaluation. There are a range of performance measures that can be used to assess airport security screening technologies, including those focused specifically on the technology and those specific to the human operator.

This section provides an overview of the key measures of performance for evaluating the role of the human operator.

5.2.1 Detection Accuracy

Perhaps the most important measure of performance is detection accuracy in terms of what proportion of persons of interest or items of contraband are detected, and how many false alarms occur. In addition, the accuracy of the human operator may also be defined in terms of more specific measures derived from the Signal Detection Theory (SDT) framework [84].

When the image of a person of interest or an item of contraband is presented to the human operator there are two possible outcomes:

1. *Hit* – correctly classifying the individual as a person of interest or the object as an item of contraband, or
2. *Miss* – incorrectly failing to classify the individual as a person of interest or the object as an item of contraband.

For images of individuals who are not of interest (i.e., not on the watch list) and objects which are not contraband, two possible outcomes are also possible:

1. *False alarm* – incorrectly classifying the individual as a person of interest or the object as an item of contraband, or
2. *True reject* – correctly classifying the individual as not being a person of interest or the object as not being an item of contraband.

Generating this data would involve recording operator responses to presented face or X-ray images and noting the type of outcome (e.g., hit or false alarm) for each presented image. Details on how to then calculate these measures are detailed in [85].

However, human operator detection accuracy may not be straightforward to determine depending on the type of response the operator is required to give and the nature of the decision. For example, in the case of some imaging baggage or passenger screening technologies, the human operator might not only have to indicate that there is contraband, but also what it may consist of, where it may be located, and how it is concealed. In order to best replicate the expectations of the operational environment, a correct response may be considered to be the response that would reasonably lead to the identification of that contraband. For example, if strapping used to contain the contraband was identified, but not the contraband itself, this would be classified as a hit because it would likely lead to subsequent identification of the contraband in a physical search. This will result in error rates that best reflect what could be expected in the real-world.

5.2.2 Discrimination and Bias

SDT also provides two additional measures (discrimination and bias) that can be useful in further analysing the accuracy of human judgments. Discrimination relates to a person's ability to distinguish target and non-targets, whereas bias relates to a person's tendency to raise an alarm. Two people may have the same discrimination ability, but they may differ in their bias.

These measures are important for two reasons. Firstly, discrimination and bias are measures that describe human performance in a way that is not easily obtained by examination of only the detection performance in terms of hits and alarms.

Secondly, these measures can assist in locating the source of poor performance in airport security, (i.e., they may provide potential insight into why performance is poor and how it might be improved). This obviously has important implications for improving the security level in airports in a cost effective manner.

For example, consider the case of an individual displaying good discrimination, but high positive bias. While they are skilled at the task they are very reluctant to raise an alarm. To examine this, one might focus on why the bias is so low (i.e., are there issues surrounding business practices and culture, and undesired consequences of raising an alarm?) Alternatively, poor overall performance that results not from strong bias but from poor discrimination implicates different processes. In this case, attention should focus on problems with the operator's task-based skills, the technology that they may be using (if any), or both. Such problems may be best addressed by looking at training, employee selection, equipment calibration and performance, as well as the usability of the technology's interface.

The discrimination and bias statistics may be derived from the obtained detection accuracy performance statistics (as discussed in Section 5.2.1). Definitions for the conventional discrimination and bias measures can be found in [84], with alternative mathematical derivations of discrimination and bias, which may be more appropriate in some operational environments, in [85, 86]. Furthermore, an application of these performance measures in a human factors study of an airport imaging device can be found in Chapter 6.

5.2.3 Subjective Confidence

In conjunction with operator decisions about the status (threat/no threat) of a face match decision or X-ray image, subjective confidence, that is, the human operator's confidence in their decision can be collected. These data may be collected on a Likert scale [87], such as a five-point scale (e.g., 1 = Very Confident to 5 = Not Confident at All), although the size of the scale may be varied if desired. Alternatively, confidence ratings may be collected on a continuum scale (e.g., as a percentage). Confidence ratings should be collected immediately after the operator has made a decision.

While this confidence rating process is not present in the normal operational environment, this measure may be very useful in examining the decision making processes of the operator. For example, inaccurate responses that are associated with high confidence suggest a lack of awareness of the quality of their responses and may reflect unwarranted faith in the abilities of the technology and/or the operator's abilities. Similarly a significant positive correlation between time and confidence (i.e., that longer response times are associated with lower confidence ratings) could indicate tentativeness of response.

5.2.4 Processing Speed

If the technology itself is too slow, or does not allow the human operator to undertake their decision in a timely manner, the flow of individuals through the screening area could be disrupted. Disruptions to throughput flows can impact on various airport stakeholders, including the airlines and retailers, and may also generate negative public sentiment towards the technologies. It is often the case that specific throughput flow rates are required, and meeting these requirements may result in decreased system performance.

In terms of assessing the speed of the systems it is often important to consider the time taken by both the technology and the human operator to make a decision. The mean time and standard deviation should be considered[2] for both the human and technical parts of the system individually, as well as the system overall. During the analysis any outliers should be examined, especially if artefacts of the trial process may have contributed to large variations in the time taken.

To assist in the assessment of processing speed of the system, it is useful to video record both the movement of individuals and objects through the technologies and also the responses of the human operators. It is important that the video feeds are time synchronised and recorded directly onto DVD disks or hard drive for later analysis. The collection of processing times involves defining appropriate start and stop times (and associated cues for determining these) and noting the relevant time on the basis of the time date stamp on the video recordings. The procedure for the collection of process timings is discussed in further detail in Chapter 4.

5.2.5 Cognitive, Perceptual and Other Abilities

As discussed in Section 4.5, it is likely that particular individual abilities may be associated with better performance on human operator tasks associated with security screening technologies. It may be that some individuals are more naturally suited to security screening task of this kind than others. Alternatively, it may be that some individuals perform well only on a particular type of security screening technology, or even only on a particular machine within a security screening technology class, but not so well on others.

As part of evaluations, tests of cognitive, perceptive and other abilities that may be of relevance (e.g., personality) could be administered to operators. An example of this approach can be found in Chapter 6, while details of appropriate tests can be found in [53, 60, 88, 89, 90, 91, 92, 93]). The results obtained from these tests could then be related to the operator's detection performance, discrimination and bias measures, speed and confidence ratings. This may provide a useful explanation for the results obtained for these measures. Additionally, these results may potentially be used as a basis for identifying the types of individuals that may be best suited as operators on the security screening technologies and/or equipment under evaluation.

5.2.6 Usability Assessment

Operator-based usability assessments of security screening technologies may provide insights into factors, not previously considered, that affect overall system performance. The identification of such factors will hopefully result in improvements to the technology and/or business practices. These factors can be assessed using a variety of methods as outlined below and in [94].

---

[2] Assuming that the distribution of times obtained is normal.

5.2.6.1 Qualitative Observations

During an evaluation, qualitative observations of operators interacting with the system should be undertaken. This may highlight common errors that are experienced in using the system, from which frequency observations could be collected to generate statistics on the prevalence of these issues. An understanding of the common issues experienced by operators may suggest ways in which the system could be modified to improve performance. These results could be linked to processing speed (as was discussed in Section 5.2.4) to provide an indication of how errors affect processing time and functionality. The procedure for the collection of qualitative observations is discussed in further detail in Chapter 4.

5.2.6.2 Questionnaires

Usability questionnaires may also be used to complement findings obtained from qualitative observations (as discussed in Section 5.2.6.1). A usability questionnaire (for examples see [95, 96, 97]) is designed to solicit the human operator's view on each of the elements of the system under examination, covering both the performance and usability of the system.

This type of questionnaire often consists of multiple parts, including optional demographic questions (e.g., details of the operator's current employment, training and experience) and a series of performance and usability questions. The performance and usability section of the questionnaire consists of system related questions, typically in terms of the operator's perception of the effectiveness and efficiency of the system, and their overall satisfaction with using the system. The human operators are often asked to respond on a Likert scale [87] about various aspects of the system, including rating themselves as users on the specific system being examined. Finally, the operators may be asked to list, in a free-text format, the things that they liked and disliked about the specific system being examined.

If deemed necessary, follow-up questionnaires can be given to the operators which target specific issues uncovered in earlier questionnaires.

5.2.6.3 Interviews and Focus Groups

In order to obtain more in-depth information about usability and performance issues in the specific operational context, focus groups and/or interviews may be conducted with operators. It is necessary to conduct interviews with a number of different users to obtain a broad range of views from that operational setting.

Interviews should be flexible and consist of open-ended questions designed to elicit a descriptive narrative of the procedures and practices used. Additionally, opinions of the operators on the functionality of the current system, and ways in which it may be modified and improved, could be discussed.

6 Challenges in Evaluating Human Operator Performance

There are numerous challenges in evaluating the performance, and in particular, the feasibility, of security screening systems. Some of these challenges are due to the evaluation type and/or airport environment, as a consequence of the systems themselves, or the desire to produce scientifically valid and relevant results that can form the basis for an informed decision on the likely impact of the implementation of a system in the airport environment.

The following sections focus on the challenges which are related to the role of the human operator.

6.1 Methodology Design

As discussed in Section 5.1, there are several different types of evaluations which may be utilised in the multi-phase evaluation of a system. The technical assessments will provide limited guidance as to the relative merits of comparable systems in terms of the role of the human operator. Similarly, scenario tests are not entirely realistic, since they cannot simulate the complex airport environment and, in particular, the various factors which will impact on the performance of the human operator (as discussed in Section 4).

There are numerous challenges in conducting operational trials in airports. In terms of assessing the role of the human operator, there may be limited opportunities to collect data and difficulties in observing interactions due to the requirements to minimise the impact on the existing business processes or because of concerns with evaluating the performance of staff members undertaking a 'security' function.

In order to collect sufficient data (see Section 6.2), one approach can be to conduct scenario tests in the airport environment, often in off-peak times or when no members of the public are present. However, it is often difficult to assess the relevance of this data, since many of the factors which can impact on human operator performance (i.e., fatigue and workload) will either not be relevant, or will be different than would be the case during peak times. In addition, due to the need to collect a sufficient amount of data to achieve sufficient confidence in the resulting statistics (see Section 6.2), it is usually necessary to ensure that events of interest occur far more frequently than would be the case for a system which was deployed. Human operators (unlike the technologies), however, have the capacity to remember previous events, which limits the ability to use subjects or items multiple times. This may have an impact on the human operator's decision process.

For face recognition systems and passenger screening systems which require a privacy filter, thresholds are required to filter the images that are provided to the human operator. If these thresholds are not set appropriately the resulting data can be of little value. For example, if the threshold for a face recognition system is too high/low then the number of alarms presented to the human operator will be less/more than anticipated which will impact on the human operator's performance. Due to this issue an assessment of the 'baseline' performance (i.e., matching performance across multiple thresholds) of the face recognition system is essential prior to incorporating human operators in the trial.

6.2 Sample Size Considerations

When testing security screening systems, using either scenario or operational methods, there are good statistical reasons to obtain the maximum possible amount of data, and to maximise the diversity of that data. If the number of tests is too small, then individual outlier results can significantly bias the measured outcomes. For example, one particular human operator in a test may be unable to properly interpret imagery from an X-ray imager. If this operator is the only one used for testing such a system, the results may be much worse than would be obtained by a more competent operator. Thus, there is a need to include testing of as many operators as possible and also to use as many test subjects/items as possible to obtain the most reliable estimate of performance.

Task completion theory has been developed to allow an estimate to be made of the reliability of data obtained from small samples [98]. This theory is based on an assumption that all variability in human operators or test subjects/items is normally distributed and that all measurements are independent and randomly extracted from an overall distribution. These assumptions may not strictly hold in any test, but the methods still allow an estimate of the reliability of data to be made.

The basic question that is answered by the sampling theory is that given the limited number of tests, how confident can a researcher be that the results obtained are close to the results that would be obtained from an unlimited number of tests? For example, the theory indicates that, if the probability of a hit is measured to be 60% from 100 tests, then we can be 95% confident that the underlying probability of a hit will be somewhere in the region between 50% and 69%. With 1000 tests, and 60% probability, the 95% confidence bounds would be 57%-63%, so increasing the number of tests can significantly improve the accuracy of an estimate.

In practice, the number of trained human operators that can be tested will be low, because the numbers of such people are generally limited and they are in demand for other purposes. Thus, it is a good idea to plan to use as many human operators as possible, but to be aware that sampling issues may limit the validity of the results. It is sometimes possible to use a larger number of test subjects, either by calling for volunteers or by recruiting paid subjects. Also, it may be possible to re-use test subjects/items to increase the number of tests, but, although this is a widely employed technique [82], it should be used sparingly, since it can introduce bias into the results because the tests are no longer independent.

As a general rule of thumb, based on past experience, a minimum of at least 50 different test subjects (and preferably more than 100) are required for scenario or operational trials, with test subjects used a maximum of five times in any test series to minimise bias and to minimise any build-up of human operator familiarity. For testing which involves items to be detected that are either fitted to test subjects or placed in baggage, the position of the items should be changed for each test sequence to increase the variability.

In some applications (e.g., face recognition), a watch list is employed and matches will be made with that list. Some detection/matching systems employ software to aid the operator by keeping the false alarm rate constant and therefore, as a consequence, the results from testing with one watch list size cannot be readily extrapolated to other watch list sizes. Extrapolation is also not possible where there is any filtering of which algorithm matches are presented to the human operator (e.g., only the top few matches for each acquired image) or where watch list image quality varies. Thus, it is necessary to carefully assess the watch list requirements before testing to ensure that the results will be applicable in the planned operational environment. The significance of the watch list issues often do not become clear until after testing has been carried out and the data analysis is underway, so in some circumstances it may be worthwhile to carry out some initial testing to make sure that the bulk of the tests are useful.

6.3 Context of the Results

Another challenge of evaluating security screening systems is ensuring that the results are not only presented in the right context, but that they also address the fundamental questions of the client. This is particularly relevant for systems incorporating human operators, since the results could be used not only to consider the feasibility/suitability of the systems, but also the impact of the human operator.

For the client, the decision to implement a security screening system is about the costs and benefits of the system. The costs are not only financial but can also be, for example, decreased throughput rates, false alarms and negative public opinion. Whilst the benefits of such systems are clear (enabling threat detection and response, as well as providing a deterrent effect) assigning a value to these is difficult, especially given that, in most cases, the likelihood of a threat may be low.

Often the focus of systems such as those considered here is on the error rate; that is, the system failed to identify a percentage of persons of interest or items of contraband. Whilst it would obviously be better to achieve an error rate of 0%, given the trade-off between threats and false alarms this is not possible in an operational airport environment with the current systems. Instead the focus should be on the performance of the system in comparison to the ‘baseline’ (i.e., whatever the current capability is).

7 Conclusions

The role of the human operator within security screening technologies discussed in this paper is currently (and is likely to remain) a critical component in moderating the performance of all of these systems. Although security screening technologies have some degree of automation, all such systems still rely on a human operator to make the final determination of whether a person of interest or an item of contraband has been detected. Nevertheless, human operator performance (and therefore the performance of these systems) may be compromised by failing to account for human fallibility, or by failing to provide an optimal working environment (e.g., technology, physical environment, business processes etc).

Despite this, DSTO’s experience in evaluating the performance of these types of systems is that very little consideration is given to the role of the human operator in the design of the technologies and also when implementing the systems in the operational airport environment. Given the importance of the role of the human operator in the screening technologies discussed here, in that such systems cannot achieve useful levels of performance in the operational environment without human operators, it is critical that all evaluations consider the role of the operator in the assessment of the feasibility of screening technologies.

References

1. Kaine, A.K.: The impact of facial recognition systems on business practices within an operational setting. In: Proceedings of EC-VIP, 4th EURASIP Conference on Video/Image Processing and Multimedia Communications, vol. 2, pp. 315–320 (2003), `http://www.vcl.fer.hr/ec2003/programme.pdf`
2. Butavicius, M.A.: Evaluating and predicting the performance of an identification face recognition system in an operational setting. Australian Society for Operations Research Bulletin 25(2), 2–13 (2006)
3. Sunde, J., Butavicius, M.A., Graves, I., Hemming, D., Ivancevic, V., Johnson, R., Kaine, A., McLindin, B.A., Meaney, K.A.: Methodology for evaluating the operational effectiveness of facial recognition systems. In: Proceedings of EC-VIP, 4th EURASIP Conference on Video/Image Processing and Multimedia Communications, vol. 2, pp. 441–448 (2003), `http://www.vcl.fer.hr/ec2003/programme.pdf`
4. Shea, D.A., Morgan, D.: Detection of explosives on airline passengers: Recommendation of the 9/11 Commission and Related Issues. GAO-07-634 Aviation Security paper (2007)
5. `http://www.customs.gov.au/site/page.cfm?u=5552` (last accessed April 2009)
6. Hannum, D.W., Parmeter, J.E.: Survey of Commercially Available Explosives Detection Technologies and Equipment. National Institute of Justice 171133. The National Law Enforcement and Corrections Technology Center (1998)
7. Parmeter, J.E., Murray, D.W., Hannum, D.W.: Guide for the Selection of Drug Detectors for Law Enforcement Applications. National Institute of Justice Guide 601- 00. Law Enforcement and Corrections Standards and Testing Program (2000)
8. Singh, S., Singh, M.: Explosives detection systems (EDS) for aviation security. Signal Processing 83, 31–55 (2003)
9. Harel, A.: Biometrics, Identification and Practical Ethics (2001), `http://www.europeanbiometrics.info/images/resources/117_934_file.doc` (last accessed August 19, 2008)
10. Keval, H., Sasse, M.A.: Can we ID from CCTV? Image quality in digital CCTV and face identification performance. In: Proceedings of the SPIE Conference on Mobile Multimedia/Image Processing, Security and Applications, 69820K-69820K-15 (2008)
11. Sheridan, T.B.: Humans and automation: Systems design and research issues. Human Factors and Ergonomics Society/Wiley, Santa Monica/New York (2002)
12. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse and abuse. Human Factors 39, 230–253 (1997)
13. Wickens, C.D., Lee, J.D., Liu, Y., Gordon Becker, S.E.: An Introduction to Human Factors Engineering, 2nd edn. Pearson, Upper Saddle River (2004)
14. Parasuraman, R.: Human-computer monitoring. Human Factors 29, 695–706 (1987)

15. Vast, R.L., Butavicius, M.A.: A literature review of face recognition for access control: Human versus machine solutions. In: DSTO-TR-1747, Defence Science and Technology Organisation, Edinburgh, SA (2005)
16. Pike, G., Kemp, R., Brace, N.: The Psychology of Human Face Recognition. Paper presented at the IEE Electronics and Communications: Visual Biometrics, London (2000)
17. Burton, A.M., Miller, P., Bruce, V., Hancock, P.J.B., Henderson, Z.: Human and automatic face recognition: a comparison across image formats. Vision Research 41, 3185–3195 (2001)
18. Hancock, P.J.B., Bruce, V., Burton, A.M.: Recognition of unfamiliar faces. Trends in Cognitive Sciences 4(9), 330–337 (2000)
19. Lablack, A., Djeraba, C.: Analysis of human behaviour in front of a target scene. In: 19th International Conference on Pattern Recognition, ICPR 2008 (2008)
20. Harrigan, J.A., Taing, K.T.: Fooled by a smile: detecting anxiety in others. Journal of Nonverbal Behavior 21(3), 203–221 (1997)
21. Harrigan, J.A., Harrigan, K.M., Sale, B.A., Rosenthal, R.: Detecting anxiety and defensiveness from visual and auditory cues. Journal of Personality 64(3), 675–709 (1996)
22. Hill, H., Bruce, V.: The effects of lighting on the perception of facial surfaces. Journal of Experimental Psychology: Human Perception and Performance 22, 986–1004 (1996)
23. Liu, X., Gale, A.G.: Effects of time pressure on searching for terrorist threats in Xray air passenger luggage images. In: de Waard, D., Hockey, B., Nickel, P., Brookhuis, K. (eds.) Human Factors Issues in Complex System Performance. Shaker Publishing, Maastricht (2007)
24. Schneier, B.: Beyond Fear: Thinking Sensibly about Security in an Uncertain World. Copernicus Books, New York (2003)
25. Shackel, B.: Applied Ergonomics Handbook. IPC Technology Press, Guildford (1975)
26. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T.: Human- Computer Interaction. Addison-Wesley, Harlow (1994)
27. Murata, A., Hirokazu, I.: Effectiveness of Cognitively Engineered Human Interface Design. In: Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 20, pp. 2717–2720 (1998)
28. Neiderman, E.C., Fobes, J.L.: Human Factors Considerations for Aviation Security Technologies. In: Proceedings of the 33rd Annual Security Technology Conference (1999)
29. Elliot, E.S., Wills, E.J., Goldstein, A.G.: The effects of discrimination on training on the recognition of white and oriental faces. Bulletin of the Psychonomic Society 2, 71–73 (1973)
30. Goldstein, A.G., Chance, J.E.: Effects of training on Japanese face recognition: reduction of the other-race effect. Bulletin of the Psychonomic Society 23, 211–214 (1985)
31. Cesa, I.L.: Of attractive librarians and lop-sided faces: development and testing of a training procedure for improving the accuracy of eyewitnesses. Dissertation Abstracts International 56, 5221 (1996)
32. Butavicius, M., Parsons, K., McCormac, A., Johnson, R., Questiaux, D., Ayubi, Z., Sunde, J.: Scenario Testing of Body Scanners (Backscatter X-ray, Active and Passive Mm-Wave, THz), LAGs Type 3 (U), DSTO-TN-0883 (2008)

33. Schwaninger, A.: Airport security human factors: From the weakest to the strongest link in airport security screening. In: The 4th International Aviation Security Technology Symposium (2006)
34. Gerrig, R.J., Zimbardo, P.G.: Psychology and Life, 17th edn. Allyn & Bacon, Boston (2005)
35. Schwaninger, A.: Computer based training: a powerful tool to the enhancement of human factors. In: Aviation Security International, pp. 31–36 (February 2004)
36. McCarley, J.S., Kramer, A.F., Wickens, C.D., Vidoni, E.D., Boot, W.R.: Visual Skills in Airport-Security Screening. Psychological Sciences 15, 302–306 (2004)
37. Schwaninger, A., Hardmeier, D., Hofer, F.: Aviation security screeners visual abilities and visual knowledge measurement. IEEE Aerospace and Electronic Systems 20, 29–35 (2005)
38. Haga, S., Shinoda, H., Mitsuteru, K.: Effects of task difficulty and time-on-task on mental workload. Japanese Psychological Research 3, 134–143 (2002)
39. Wallach, M.A., Kogan, N., Bem, D.J.: Group Influence on Individual Risk Taking. Journal of Abnormal and Social Psychology 65, 75–86 (1962)
40. Atthowe, J.: Interpersonal decision making: The resolution of a dyadic conflict. Journal of Abnormal and Social Psychology 62, 114–119 (1961)
41. Endsley, M.R., Jones, W.M.: Situation awareness, information dominance & information warfare. AL/CF-TR-1997-0156, United States Air Force Armstrong Laboratory, Wright-Patterson Air Force Base, OH (1997)
42. Myers, D.G., Lamm, H.: The group polarization phenomenon. Psychological Bulletin 83, 602–627 (1976)
43. Menner, T., Barrett, D.J.K., Phillips, L., Donnelly, N., Cave, K.R.: Costs in Searching for Two Targets: Dividing Search Across Target Types Could Improve Airport Security Screening. Applied Cognitive Psychology 21, 915–932 (2006)
44. Parasuraman, R.: Vigilance, monitoring, and search. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) Handbook of Perception and Human Performance, vol. 2. John Wiley & Sons, Inc., Canada (1986)
45. Robert, G., Hockey, J.: Changes in screener efficiency as a function of environmental stress, fatigue, and circadian rhythms. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) Handbook of Perception and Human Performance, vol. 2, John Wiley & Sons, Inc., Canada (1986)
46. Webster, D.M., Richter, L., Kruglanski, A.W.: On leaping to conclusions when feeling tired: mental fatigue effects on impressionable primacy. Journal of Experimental Social Psychology 32, 181–195 (1996)
47. Harma, M.: Individual differences in toleration to shiftwork: A review. Ergonomics 36(1-3), 101–109 (1993)
48. Boff, K.R., Kaufman, L., Thomas, J.P. (eds.): Handbook of Perception and Human Performance, vol. 2. John Wiley & Sons, Inc, Canada (1986)
49. Ariely, D., Zakay, D.: A timely account of the role of duration in decision making. Acta Psychologica 108, 187–207 (2001)
50. Carroll-Mayer, M., Fairweather, B., Stahl, B.C.: Ignotious Per Ignotum: 21st Century Surveillance Technology and the Presumption of Guilt. In: Proceedings of the 21st British and Irish Law, Education and Technology Association Conference, Malta (April 2006)
51. Dror, I.E.: Don't forget us humans! The Times (July 31, 2006)

52. Butavicius, M., Foster, R., Whittenbury, A., Parsons, K., McCormac, A., MacLeod, V.: Human Factors Study of the ThruVision T400 Passive Terahertz Camera. In: DSTO-TN-0821, Defence Science and Technology Organisation, Edinburgh, SA (2008)
53. Hardmeier, D., Hofer, F., Schwaninger, A.: The x-ray object recognition test (XRay ORT) – A reliable and valid instrument for measuring visual abilities needed in x-ray screening. In: IEEE ICCST Proceedings, vol. 38, pp. 303–308 (2005)
54. Ellis, H.D.: Introduction to aspects of face processing: Ten questions in need of answers. In: Ellis, H., Jeeves, M., Newcombe, F., Young, A. (eds.) Aspects of Face Processing, Nijhoff, Dordrecht, The Netherlands, pp. 3–13 (1986)
55. Gauthier, I., Behrmann, M., Tarr, M.J.: Can face recognition really be dissociated from object recognition? Journal of Cognitive Neuroscience 11, 349–370 (1999)
56. Gauthier, I., Logothetis, N.K.: Is face recognition so unique after All? Journal of Cognitive Neuropsychology 17, 125–142 (2000)
57. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Surveys 35, 399–458 (2003)
58. Goren, C., Sarty, M.E., Wu, P.Y.K.: Visual following and pattern discrimination of face-like stimuli by newborn infants. Pediatrics 56, 544–549 (1975)
59. Slater, A., Quinn, P.C.: Face recognition in the newborn infant. Infant and Child Development 10, 21–24 (2001)
60. Pacini, R., Epstein, S.: The relation of rational and experiential information processing styles to personality, basic beliefs and the ratio-bias phenomenon. Journal of Personality and Social Psychology 76(6), 972–987 (1999)
61. Furl, N., Phillips, P.J., O'Toole, A.J.: Face recognition algorithms and the other race effect: Computational mechanisms for a developmental contact hypothesis. Cognitive Science 26, 797–815 (2002)
62. Shapiro, P.N., Penrod, S.: Meta-analysis of facial identification studies. Psychological Bulletin 100, 139–156 (1986)
63. Malpass, R.S., Lavigneur, H., Weldon, D.E.: Verbal and visual training in face recognition. Perception and Psychophysics 14, 285–292 (1973)
64. Wild, H.A., Barrett, S.E., Spence, M.J., O'Toole, A.J., Cheng, Y.D., Brooke, J.: Recognition and sex categorization of adults' and children's faces: Examining performance in the absence of sex- stereotyped cues. Journal of Experimental Child Psychology 77, 269–291 (2000)
65. Wright, D.B., Sladden, B.: An own gender bias and the importance of hair in face recognition. Acta Psychologica 114, 101–114 (2003)
66. de Frias, C.M., Nilsson, L.-G., Herlitz, A.: Sex differences in cognition are stable over a 10-year period in adulthood and old age. Aging, Neuropsychology, and Cognition 13, 574–587 (2006)
67. Rehnman, J.: The role of gender in face recognition. Department of Psychology, Stockholm University (2007)
68. Rehnman, J., Herlitz, A.: Women remember more faces than men do. Acta Psychologica 124(3), 344–355 (2007)
69. McKelvie, S.J., Standing, L., St. Jean, D., Law, J.: Gender differences in recognition memory for faces and cars: Evidence for the interest hypothesis. Bulletin of the Psychonomic Society 31, 447–448 (1993)
70. Cellerino, A., Borghetti, D., Sartucci, F.: Sex differences in face gender recognition in humans. Brain Res. Bull. 63(6), 443–449 (2004)

71. Lewin, C., Herlitz, A.: Sex differences in face recognition- women's faces make the difference. Brain & Cognition 50, 121–128 (2002)
72. Shepherd, J.: Social factors in face recognition. In: Davies, G., Ellis, H., Shepherd, J. (eds.) Perceiving and Remembering Faces, pp. 55–79. Academic Press, San Diego (1981)
73. Slone, A.E., Brigham, J.C., Meissner, C.A.: Social and cognitive factors affecting the own-race bias in whites. Basic and applied social Psychology 22(2), 71–84 (2000)
74. Wright, D.B., Stroud, J.S.: Age differences in lineup identification accuracy: People are better with their own age. Law & Human Behavior 26, 641–654 (2002)
75. Anastasi, J.S., Rhodes, M.G.: Evidence for an own-age bias in face recognition. North American Journal of Psychology 8, 237–253 (2006)
76. Dror, I.E., Fraser-Mackenzie, P.A.F.: Cognitive Biases in Human Perception, Judgment and Decision Making: Bridging Theory and the Real World. In: Rossmo, K. (ed.) Criminal Investigative Failures. Taylor & Francis, London (2008)
77. Bright, D.A., Goodman-Delahunty, J.: The influence of gruesome verbal evidence on mock juror verdicts. Psychiatry, Psychology and Law 11, 154–166 (2004)
78. Dror, I.E., Peron, A.E., Hind, S., Charlton, D.: When Emotions Get the Better of Us: The Effect of Contextual Top-down Processing on Matching Fingerprints. Applied Cognitive Psychology 19, 799–809 (2005)
79. Stacey, R.B.: Report on the erroneous fingerprint individualization in the Madrid train bombing case. Journal of Forensic Identification 54, 706–718 (2004)
80. Gilovich, T., Griffin, D., Kahneman, D.: Heuristics and biases: The psychology of intuitive Judgment. Cambridge University Press, New York (2002)
81. Ruthruff, E.: A test of the deadline model for speed–accuracy tradeoffs. Perception & Psychophysics 58(1), 56–64 (1996)
82. Mansfield, A.J., Wayman, J.L.: Best practices in testing and reporting performance of biometric devices Version 2.01, National Physics laboratory NPL Report CMSC (February 14, 2002)
83. Phillips, P.J., Martin, A., Wilson, C.L., Przbocki, M.: An introduction to evaluating biometric systems. Computer 33(2), 56–63 (2008)
84. Green, D.M., Swets, J.A.: Signal Detection Theory and Psychophysics. Wiley, New York (1966)
85. Stanislaw, H., Todorov, N.: Calculation of signal detection theory measures. Behavior Research Methods, Instruments & Computers 31(1), 137–149 (1999)
86. Butavicius, M.A.: Evaluating and predicting the performance of an identification face recognition system in an operational setting. Australian Society for Operations Research Bulletin 25(2), 2–13 (2006)
87. Likert, R.: A Technique for the Measurement of Attitudes. Archives of Psychology 140, 1–55 (1932)
88. Harris, D.H.: How to Really Improve Airport Security. Ergonomics in Design, 1–6 (Winter 2002)
89. Leyk, D., Sievert, A., Heiss, A., Gorges, W., Ridder, D., Alexander, T., Wunderlich, M., Ruther, T.: Validation of a short-term memory test for the recognition of people and faces. Ergonomics 51, 1125–1136 (2008)
90. Megreya, A.M., Burton, A.M.: Unfamiliar faces are not faces: Evidence from a matching task. Memory & Cognition 34, 865–876 (2006)
91. Cohen, R.J., Swerdlik, M.E.: Psychological Testing and Assessment, 6th edn. McGraw-Hill, New York (2005)

92. Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M., Sagi, D.: Perceptual learning in contrast discrimination: The effect of contrast uncertainty. Journal of Vision 4(12), 993–1005 (2004)
93. Goldstein, E.B.: Sensation and Perception, 6th edn. Wadsworth, California (2002)
94. Pew, R.W., Mavor, A.S. (eds.): Human-System Integration in the System Development Process: A New Look. National Academies Press, Washington (2007)
95. Jordan, P.W.: Usability Evaluation in Industry. CRC Press, Boca Raton (1996)
96. Dumas, J.S., Loring, B.A.: Moderating Usability Tests: Principles and Practices for Interacting. Morgan Kaufmann, New York (2008)
97. Rubin, J., Chisnell, D.: Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd edn. Wiley, New York (2008)
98. Sauro, J., Lewis, J.R.: Estimating completion rates from small samples using binomial confidence intervals: Comparisons and recommendations. In: Proceedings of the Human Factors and Ergonomics Society, 49th Annual Meeting, pp. 2100–2104 (2005)

Chapter 6

Assessment of the ThruVision T4000 Passive Terahertz Camera: A Human Factors Case Study

M.A. Butavicius, K.M. Parsons, A. McCormac, R. Foster,
A. Whittenbury, and V. MacLeod

Defence Science and Technology Organisation, Edinburgh, Australia

Abstract. This chapter describes an experiment conducted by Australia's Defence Science and Technology Organisation (DSTO) on the ThruVision T4000 Passive Terahertz (THz) camera. The purpose of this study was to assess a number of the human factors issues surrounding the use of this device in an airport scenario. It presents the theoretical background, methodology (including psychological pre-tests) as well as aspects of the results from a controlled psychological experiment.

1 Background

In airports today, metal detectors are traditionally used to detect metallic contraband items concealed under a person's clothing. However, passive THz imagers, such as the Thruvision T4000 Passive THz camera, have the ability to detect both metallic and non-metallic concealed objects and, by providing an image of the THz waves, they can also provide information on the size, shape and location of the contraband. Passive THz imagers work by detecting the natural THz radiation from a scene. In the context of passenger screening, their success is based on the assumption that contraband items such as weapons and drugs will have reflectivity and emissivity distinguishable from human skin. Because THz radiation can penetrate fabrics and plastics, the device can effectively see through the clothes of a passenger to potentially reveal concealed items.

Unlike many screening devices where the user simply responds to or verifies an alarm that is raised by the device, there is no such automation in the THz imaging device and software. In addition, the angular resolution of the image is low and previous research at DSTO has demonstrated that factors such as the material type and thickness of clothing as well as characteristics of the passenger's body can significantly alter the visibility of contraband. As a result, the overall performance of the system (i.e., the device and the operator working together) is highly reliant on the capabilities of the human to correctly interpret the images. In other words, the actions of the user have a vital role in the usefulness of such devices in the operational context.

In processing and acting on the information presented on the device's screen, the human is relying on their perceptual and cognitive abilities. As has been demonstrated in Vickers, Butavicius, Lee and Medvedev (2001), the human brain is capable of quickly and effortlessly analysing vast amounts of visually presented

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 183–206.
springerlink.com

information to find solutions to problems that are computationally intractable, i.e., problems for which a guaranteed solution cannot be calculated by a computer in a realistic timeframe. Given the strength of the human visual capacities, it is important to design passenger imaging devices that complement the strengths of the human visual system and to explore the elements of human information processing that may improve performance on imaging tasks.

Conventionally, the analysis of airport security devices occurs in three major stages consisting of technical, scenario and operational evaluations (for further detail see Butavicius, 2006). However, in some circumstances it is necessary to conduct additional testing that falls outside these definitions. The current study is a special case because it focuses on aspects of the user rather than just overall performance of the system. To this end, the methodology called for a large number of participants to act as screeners as well as the inclusion of additional psychological tests that would not be found in traditional technical, scenario or operational evaluations.

1.1 Previous DSTO Study

In Hall, Bird, Johnson, Resnyansky and Sunde (2007), a scenario evaluation of the ThruVision T4000 camera was conducted. In this study, the THz camera was set up in an airport in a passenger flow area. A small number of airport screeners, working singly or in pairs, viewed live footage of the THz device alongside CCTV footage of the same scene. This footage was viewed in a room separate from the passengers. During the trial, a number of staff from the Australian Customs and Border Protection Service walked past the cameras and some of these staff wore concealed items under their clothing. The staff walked past the camera multiple times and each time all of the individuals were imaged by the device this was considered a separate ‘run’. The task of the screeners was to identify the staff members who were carrying concealed items in the trial. This identification was communicated via radio to ground officers by relaying the unique number tags worn by suspects as viewed from the accompanying CCTV footage.

In this study a number of observations were made both during the trial and in a post-hoc analysis of video footage of the screeners, regarding human factors issues that may have been affecting performance. Firstly, it was noted that there were potential differences between cases where one operator was present and two operators were present. A priori, two operators might seem advantageous because they could split the workload across the various tasks. Reporting a passenger carrying contraband was done via radio and the operator needed to match the CCTV footage of the passenger with their THz image in order to identify the unique numbers on their clothing which were worn for the purposes of the trial. In addition, the flow of passengers could not be halted by an operator which meant that this reporting of a suspected contraband occurred while still monitoring new passengers. Similarly, a single operator could miss an anomaly in a THz image that a second operator could spot. In addition, in the runs with two operators there was often significant discussion between the individuals before a decision was made about whether an individual was carrying contraband. As will be discussed in more detail in section 2.3.1, this interaction may or may not necessarily be beneficial depending on the types of personalities involved. However, because of the small number of runs and the use of different mock ‘passengers’ and contraband across the runs, any valid statistical comparison between runs with one operator versus those with two was not possible. This observation led to the need for a controlled human factors trial to determine whether the use of two operators, and their interaction, would result in better performance in the airport scenario as opposed to just one operator. This issue has important practical implications for operating these devices in the airport environment because using two operators instead of one would significantly increase operating costs for this device and such a measure would need to be justified with sound evidence of a significant performance improvement.

Secondly, there was significant variation between the runs which seemed associated with individual differences, i.e., there was variation in performance that seemed strongly linked to the different operators. However, because operators saw different ‘passengers’ carrying different contraband in each run, it was not possible to determine the extent to which some operators actually outperformed others.

The question was raised as to whether some screeners were simply better at the task of interpreting the THz imager than others. In addition, if such variation is present, what individual attributes are associated with better performance and can tests for these attributes be used to select the best individuals for this role?

Thirdly, it was unknown what sort of training should be provided to screeners to operate this device. On the one hand, dedicated training programs that provide experience with a range of different cases where passengers do and do not carry contraband might better prepare them for the real life scenario. On the other hand, on the job training is a far cheaper method of training screeners in the use of this device. However it has the downside that trainees only get feedback on their correctness of their decision when they do raise an alarm and not for cases when they do not. In other words, they will not receive any information on the accuracy of their decisions when they do not raise an alarm. This limitation on the feedback provided to the screeners may, in turn, reduce their performance using the THz imager by creating a bias in their decision making.

1.2 Previous Research Outside DSTO

There were no published examples of trials or studies with a THz imager. In addition, there was little research into human performance, personality variables and cognitive and visual processing in the context of airport security. There is some research into human performance involving visual tasks in airport scenarios. This involved face matching for access control in controlled settings both in the laboratory (Lee, Vast & Butavicius, 2006; Fletcher, Butavicius & Lee, 2008) and in an airport scenario (Butavicius et al., 2008). However, the task of matching images is different from identifying images and, unlike the THz imager, this previous research also relied on images from the visible spectrum.

Previous studies involving human performance in airport security settings have mostly been concerned with the use of X-ray technologies to detect metallic contraband in luggage (Schwaninger, 2004; McCarley et al., 2004; Schwaninger, Hardmeier & Hofer, 2005). Threat Image Projection (TIP) technology has been developed as a means to test operators and to keep them alert while scanning X-ray images (Schwaninger, 2004). The TIP system projects images of threat items onto X-ray images of passenger bags at random intervals during an operator's shift (Schwaninger, 2004). This form of 'testing' allows for continued monitoring of the operators' effectiveness in the detection of threats, as well as providing more frequent 'alarm' rates than would be expected during normal operational screening.

Schwaninger, Hardmeier and Hofer (2005) also examined the efficacy of certain tests for the selection of X-ray screening personnel. One such test is the Prohibited Items Test (PIT), in which different items from an international prohibited items lists are placed into passenger bags in such a way that the item is clearly visible. Any missed detections can be assumed to be a result of a lack of visual knowledge of the threat items on the part of the operator. This knowledge is based on training and expertise (Schwaninger, 2004). Schwaninger et al. (2005) also developed the X-Ray Object Recognition Test (X-Ray ORT) which examines the image-based factors associated with x-ray baggage screening. Hardmeier, Hofer and Schwaninger (2005) have demonstrated in controlled psychological experiments that this is a valid and reliable tool and, as such, it may be beneficial as a pre-assessment tool or competency check for screeners.

Specifically, the X-Ray ORT measures three different factors: Viewpoint, Superposition and Bag Complexity. These three variables mirror the challenges faced by the screener in the real-world task. Viewpoint refers to the angle at which the threat object is viewed. Previous psychological experiments in visual object identification have demonstrated that for objects of a particular physical structure, viewing them from an angle such that their main axis is foreshortened can significantly impact identification performance (e.g., Humphrey & Jolicoeur, 1993). In the context of X-ray analysis of baggage, an example of this is viewing a gun from an angle as if the viewer was looking directly down the barrel. From this perspective, many of the characteristic features of the gun's shape are occluded resulting in difficulties in the later stages of object identification.

The second variable in the X-Ray ORT, Superposition, refers to cases where non-threat items partially obscure the image of the threat item in the bag making the target less visible. In many theories of object recognition such as Marr's (1982) computational theory of vision and Biederman's (1987) *Recognition-by-Components* theory, both Viewpoint and Superposition hinder the process of object recognition in the higher stages of visual processing. The third variable, known as Bag Complexity, refers to the amount of distractor or non-target items present in bag. In other words, Bag Complexity increases as the number of non-threat items increases. This mirrors research in visual processing where the impact of such clutter on the speed of identification has been previously identified in similar experiments (known as visual search tasks) on the detection of simple shapes (Palmer & McLean, 1995). All three variables represent visual abilities required by X-ray operators.

The X-Ray ORT represents a successful application of psychological theory to the practical problem of screening in airports. This chapter describes a similar application of psychological theory and experimentation to the THz imaging device. Our study is different in two main ways from the research investigating the X-Ray ORT. Firstly, the visual task of the THz imager is different from that of the

X-ray baggage screener because the screener will not normally be able to identify the object itself but, at best, will only be able to indicate the approximate size, shape and position of the item due to the poor resolution of camera. In DSTO testing even the background shape of the body can be distorted significantly. This reduces the human task to one of identifying any abnormalities that appear as discolourations or “holes” in the THz image. An example of such a THz image, with contraband concealed around the waist is shown in Figure 1. As a result, we will use simple visual tasks to explore the visual abilities used in this task. Secondly, in addition to visual processing abilities, we will also examine the influence of personality and cognitive processing styles on human performance on the THz imager. These factors could potentially play a significant role in influencing the decisions of the screeners in addition to their raw visual processing abilities.

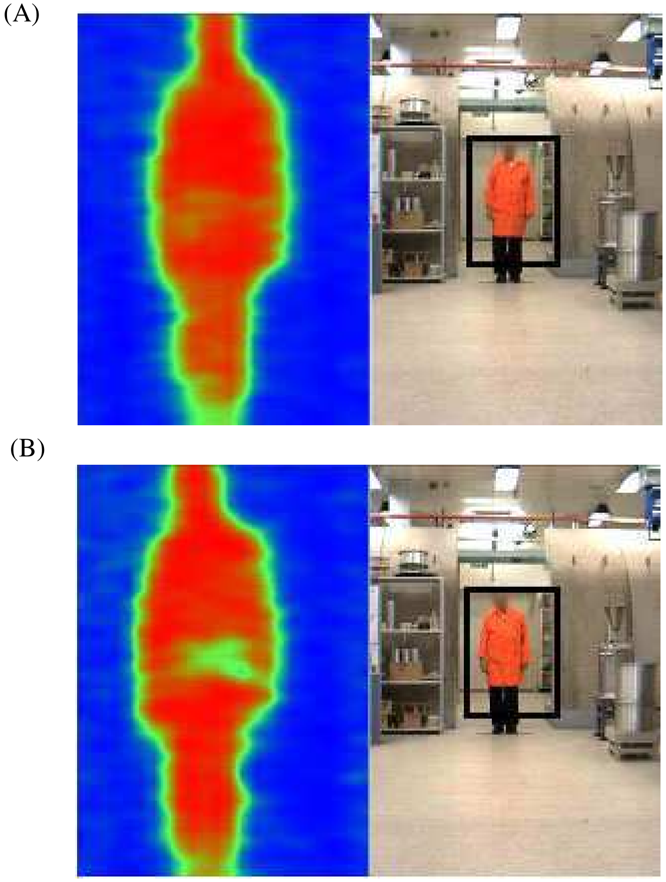


Fig. 1. Two screenshots of examples of images produced by the THz viewer. (A) shows an example of an individual without contraband while (B) depicts an individual carrying contraband around their waist concealed under their clothing. The THz image is shown to the left of the picture while the screen shot of the live camera feed is shown on the right with the corresponding area viewed by the THz camera indicated by the rectangle.

In the following section we will present an overview of the psychological theory that was brought to bear on our study of the THz imager. This includes discussion of group effects, the influence of feedback and learning, visual processing abilities and cognitive processing style. This is then followed by the aims of the experiment, a description of the experiment design and then an overview of the results of this study.

1.3 Theoretical Background

1.3.1 Group Interaction

As mentioned previously, an analysis of the footage from Hall et al. (2007) indicated that, when operators were working in pairs, there appeared to be factors present that were not at hand when operators worked alone. These effects, known as group interactions, may have influenced performance on the task. It is therefore important to examine the performance of individuals versus groups, to determine whether these factors have a significant effect on the ability to successfully operate the technology.

The area of group decision-making has been studied extensively. However, there is very little consensus regarding the differences between group and individual performance. Some research has indicated that groups tend to make more risky decisions (Wallach, Kogan & Bem, 1962), whereas other research suggests that group decisions are generally more conservative in nature (Atthowe, 1961).

Evidence suggests that poor group decision-making is often associated with poor communication skills and collaboration, and such weaknesses are particularly prevalent when pertinent information is contrary to the group norm (Endsley & Jones, 1997). Essentially, when members of a group have the same misconceptions, these incorrect assumptions can be validated by the other group members. Similarly, problems can occur when groups have shared expectations, as these expectations may limit their ability to deal with any information that does not support the preconceived idea (Endsley & Jones, 1997). In such cases, information that is contrary to expectations may be filtered out to allow the group to reach a consensus.

Related to this, teams have been found to suffer from groupthink, which is based on the desire for cohesion, and can occur when people fail to question a respected leader (Endsley & Jones, 1997). This can result in a phenomenon known as group polarisation, in which decisions tend to be more extreme than those of the individuals within the group (Myers & Lamm, 1976). There is also evidence to suggest that group decision-making can take significantly longer, particularly when decisions involve uncertainty (Atthowe, 1961). This is particularly applicable to the airport security domain, as uncertainty could be quite common, and fast and accurate decision-making is crucial in the airport environment.

In contrast, there is evidence to suggest that groups can make more effective decisions, because there is a greater chance that someone in the group will have the capability to solve the problem (Laughlin & Bitz, 1975). Effective teams are those that question the group norm and regularly self-check or reassess information (Orasanu, 1995). Furthermore, evidence suggests that individuals who are more emotionally intelligent have more effective interpersonal and social skills, and therefore, may perform more effectively within a group (Mayer, Salovey & Caruso, 2000). Essentially, group decisions are likely to be strongly influenced by the nature of the individuals within the group, as well as the role of the individual members.

1.3.2 Feedback and Learning

Feedback and learning are used in training environments to improve overall performance. Learning is a process which results in a change in behaviour, which can be measured via changes in performance (Gerrig & Zimbardo, 2005). Feedback is essentially a mechanism that facilitates learning by communicating information about performance to an individual or a group (Gerrig & Zimbardo, 2005). Therefore, when used appropriately feedback and learning can be powerful training tools.

In this study, two types of feedback were of primary interest. One type of feedback is similar to the feedback that a screener would receive while 'on-the-job' and can be best described as operational feedback. For example, in the operational environment the screener identifies a target and then receives delayed feedback that will either confirm or deny if an individual of interest was carrying contraband. In this scenario the screener receives no feedback about the passengers that they failed to identify. In other words they have no feedback pertaining to the 'missed targets', i.e., no information on who were the passengers carrying contraband that they were not able to identify. The other form of feedback is the type of feedback that can be conducted in a specific training environment. Screeners are able to receive feedback on the individuals that they believe are carrying contraband, and also receive feedback on the individuals that they fail to identify as carrying contraband. This type of feedback can be described as training feedback.

Theoretically speaking, if screeners receive only operational feedback, this may result in a bias, as they have no feedback on missed targets. For example, it is possible that the screener may unconsciously adopt a more stringent decision making technique (i.e., not raise an alarm unless they are highly certain that there is contraband present) as a way of avoiding feedback indicating they made an error. Hence, although training feedback is more time intensive, it may be important to provide proper correction of a screener's decision making during the learning process.

1.3.3 Rational Experiential Inventory

It is hypothesised that cognitive style may have an impact on operator performance (Butavicius et al., 2008). Cognitive-Experiential Self-Theory can be used to explain how different individuals process information. The theory is based on the assumption that individuals process information in a rational or an experiential mode (Epstein, 1994).

Individuals who are rational tend to make conscious and deliberate decisions; they generally exert a great deal of self control and are able to delay gratification (Pacini & Epstein, 1999). These behaviours mean that they often display low impulsivity. They therefore often demonstrate a bias towards accuracy and although they are more accurate and less error prone, they may require more processing time to reach a decision.

In contrast, experiential individuals are characterised as more natural, automatic, efficient, and spontaneous and tend to display high interpersonal skills. They generally rate very highly on impulsivity. What this means is that their visual processing is rapid, and because of this bias for speed, they are error prone as they often fail to observe the finer details (Dickman, 1985).

In an airport environment, an operator who is rational may be more motivated by accuracy. The operational setting is generally fast paced, with large volumes of passengers moving through the terminal, and the extra time that a rational individual requires means that they may not have the opportunity to identify significant numbers of suspicious targets.

In comparison, due to their rapid response time, experiential operators may be able to scrutinise more passengers using the THz images. However, they may find it more difficult to compare a THz image to a CCTV image to identify any anomalies. They may find using multiple sources of information to make an informed decision much harder than a more rational person. They may also find it challenging to recall enough details about a suspicious passenger to be able to sufficiently describe the passenger to ground staff.

The notion of rational and experiential individuals is best understood along a continuum. Although many individuals display a clear bias towards speed or a clear bias towards accuracy, others may show no bias for either rational or experiential cognitive processes. These individuals display no particular predisposition towards favouring speed or accuracy and are classified as having a medium level of impulsivity. It would appear that they may be more flexible in the way that they process visual information and they may therefore perform better on visual detection tasks than those low or high on impulsivity (Butavicius et al., 2008).

1.3.4 Visual Processing

As mentioned previously, the skill of the operator in interpreting the THz imager is likely to depend on their visual processing abilities. Therefore, simple tests of such abilities may be able to partially predict performance of an operator. If successful, such tests may be useful in pre-selection assessment for the THz operator's role.

As discussed above, the visual task itself is relatively simple – operators are required to identify when an image of a person is displaying abnormalities that may be attributable to a concealed object. These abnormalities present as "holes" or contrasting coloured blobs that feature within the outline of an individual's body and tend to be coloured similarly to the background areas of the image. To assess the individual skills that may contribute to performance on the tasks we chose to focus on three basic visual processing abilities. These were feature analysis, global precedence and signal detection.

Feature analysis tasks are based on the notion that the visual system has two sequential processing stages in producing a perception (see Treisman's (1993) *Feature Integration Theory*). In the first stage, known as the preattentive stage, the visual system detects the presence of the fundamental building blocks of perception from the image. These include basic features such as orientation, intersections, colour, movement and the ends of lines. According to Treisman's (1993) account, these features are processed in parallel, i.e., all of these components of the image across the whole display are processed at the same time. In the second stage, known as the focused attention stage, these basic features are combined to produce a more complete perception. Unlike the preattentive stage, processing in the focused attention stage is serial such that the visual system can only process a portion of the overall scene in this manner at any particular point in time.

The global precedence effect is based on our understanding of the early stages of the visual processing within the brain. The retino-geniculo-cortical system consists of two pathways: the parvocellular pathway and the magnocellular pathway. The parvocellular pathway is largely responsible for processing information from the centre of the retina (i.e., the area in the central part of what we see when we look at a scene). The information processed is concerned with fine-grain detail in the image. In comparison, the magnocellular pathway is largely concerned with processing information from the area outside of the fovea (known as peripheral vision). This pathway focuses on coarse-grain information such as overall shape and form of an image. In terms of speed of processing, the magnocellular pathway is faster than the parvocellular pathway in providing information to the higher processing areas of the brain. Because of this speed difference, the brain is able to process global information of an image faster than local aspects and this effect, known as the global precedence effect, has been confirmed in psychological experimentation (Navon, 1977; Miller, 1981).

Signal Detection Theory (SDT) is a theoretical framework that allows measurement of the sensitivity of human sensory systems (Green & Swets, 1966). In a SDT task, the observer is asked to detect the presence or absence of a signal, also known as a target, in the presence of noise. This noise comes not only from external, environmental sources but also internal noise in our sensory system which is considered to be constantly variable. The accuracy of response from both target-present and target-absent trials can then be used to measure different attributes of our visual system. In particular, the framework allows measurements of not only how well an individual can discriminate between target-present and target-absent trials but also whether they are biased towards responding "target-present" or "target-absent" in the experiment. As detailed in later sections, the SDT framework is applied to the analysis of the THz task itself (and not just the SDT visual pre-test) where the operator's task is to determine when the signal, in this case the contraband, is present in the visual image.

1.4 Summary of Research Aims

The current project extended the study of THz technology to human factors issues and was primarily concerned with testing the conditions under which operator performance improves, especially whether performance is better individually or with a partner. It also examined whether there is a relationship between performance and type of feedback received, and whether visual and cognitive assessments can predict operator performance, aiding in operator selection.

More specifically, the research aims of the project were to:

- Compare performance between one and two operator scenarios on the detection of concealed items using the THz imaging system. This will investigate whether two operators provide a significant advantage in operational performance over one operator.
- Examine whether human performance on the THz system is influenced by personality and visual processing abilities. These attributes include rational-experiential processing styles and visual perception abilities on global precedence, feature analysis and signal detection tasks. Examination of such relationships may ultimately inform the selection and screening methods for selecting operators as well as suggesting conditions for enhancing their performance.
- Compare the training benefit of a full training regime with complete performance feedback to a simple on-the-job training scenario. This will examine whether performance on the detection of contraband using the THz imager will benefit significantly from a dedicated training program.

2 Methodology

2.1 Participants

Twenty-four participants from the Defence Science and Technology Organisation took part in the study. Three of the participants were female and the average age was 35 years (SD = 14).

2.2 Materials

2.2.1 Demographic Questionnaire

A demographic questionnaire, which requested information regarding age, gender, and factors that may have the capacity to affect participant performance, including whether participants had any prior experience in screening operations, was administered at the beginning of the trial.

2.2.2 Personality and Cognitive Abilities

Pre-trial tests included a measure of rationality and experientiality (REI inventory: Pacini & Epstein, 1999). It also included three perceptual tests from Goldstein (2002). The first test of perceptual ability was a feature analysis test, in which participants were required to make judgements about whether a target item (in this case, either a cross or a circle which was green in colour) was present among a variable number of distractor items (circles and crosses in red or black). This type of test is also referred to in the literature as a visual search task. This test measured the reaction time of the participant and the accuracy of response. Figure 2 provides examples of trials with various numbers of items present.

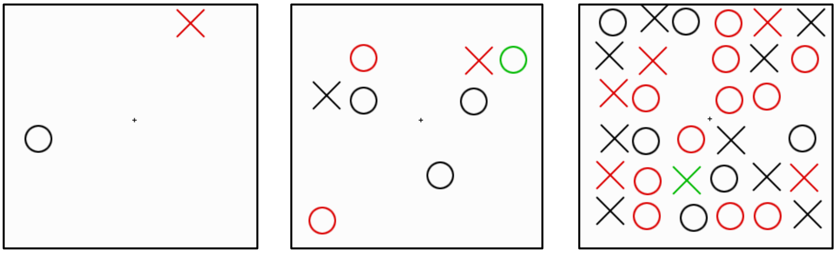


Fig. 2. Examples of feature analysis stimuli for trials with 2, 8, and 32 items, respectively. The target is present in all trials except the 2 item version.

The second perceptual ability test was a global precedence test for letters as used by Miller (1981). In this test, participants were asked to identify either the global or local aspect of an image of a large letter (either an "S" or an "H") which is made up of many smaller letters (again, either "S" or "H"). This test also returned reaction time data. Separate data was returned for each combination of letters (i.e. global S with local S, global S with local H, global H with local S, and global H with local H), as well as for trials in which the participant gave correct and incorrect responses. Figure 3 shows the four possible stimuli used in this test.

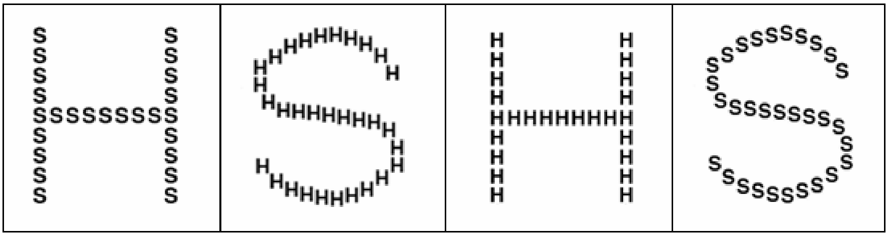


Fig. 3. Global precedence stimuli

The third test investigated the participants' signal detection ability using a variation of the Landolt (1888) broken ring task originally developed for ophthalmological testing. In this test, participants were asked to evaluate which of two circular shapes had a larger gap at the top with only a very brief presentation time. Figure 4 provides an example of the type of stimuli presented in this task. Participants were asked to provide a response to each stimulus on a 6-point rating scale where lower ratings indicated greater certainty that the wider gap in the ring was on the left and higher ratings indicated greater uncertainty that the wider gap was on the right.

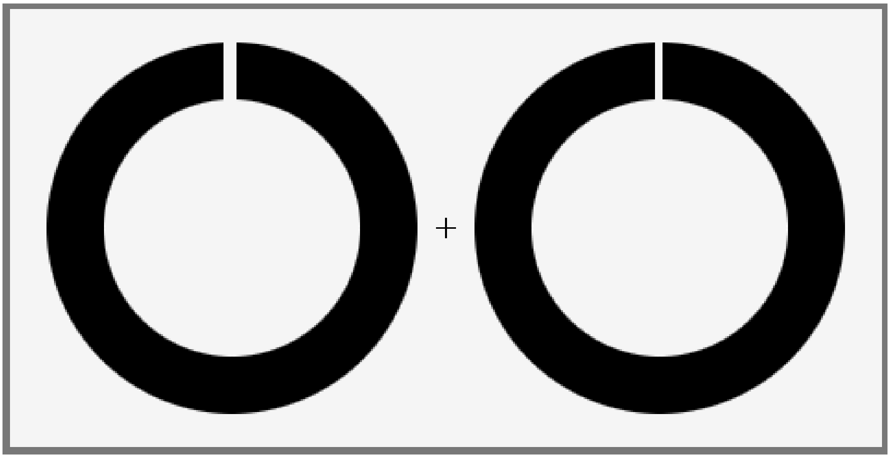


Fig. 4. Example of signal detection stimulus

2.2.3 THz Camera Footage

Participants viewed existing THz camera footage of scenario testing, including sections of four days' footage. In the previous trial, participants monitored THz camera images of individuals walking through the terminal, and indicated via radio when it appeared that a monitored individual was 'carrying' concealed contraband. The footage utilised involved continuously moving test subjects to determine the operator's ability to detect contraband in a dynamic environment. By reusing footage from scenario trial this ensured that the same stimuli were presented to each individual which controlled for differences that occur in different runs in the scenario evaluation. This in turn ensured that performance differences between individual participants were attributable to individual differences and not any variation in stimuli.

The experimenter recorded the decisions made by each participant. This included the identification number of the 'passenger' and the location of the suspected item. The experiment was also recorded using a video camera.

2.3 Procedure

Before the commencement of the experiment, participants were given details of their rights, and a description of the study, including their role if they chose to take part. Once participants gave their written consent, they were asked to complete the short demographic questionnaire. This was followed by the personality and visual tests and then the THz experiment.

As depicted in Figure 5, the experiment was divided into multiple sessions, conducted over four consecutive days. Two of these days were used to collect demographic and other information, and to provide training in the operation of the software, and then experimental data was collected over the following two days. The participants were randomly allocated to the treatment groups, such that there

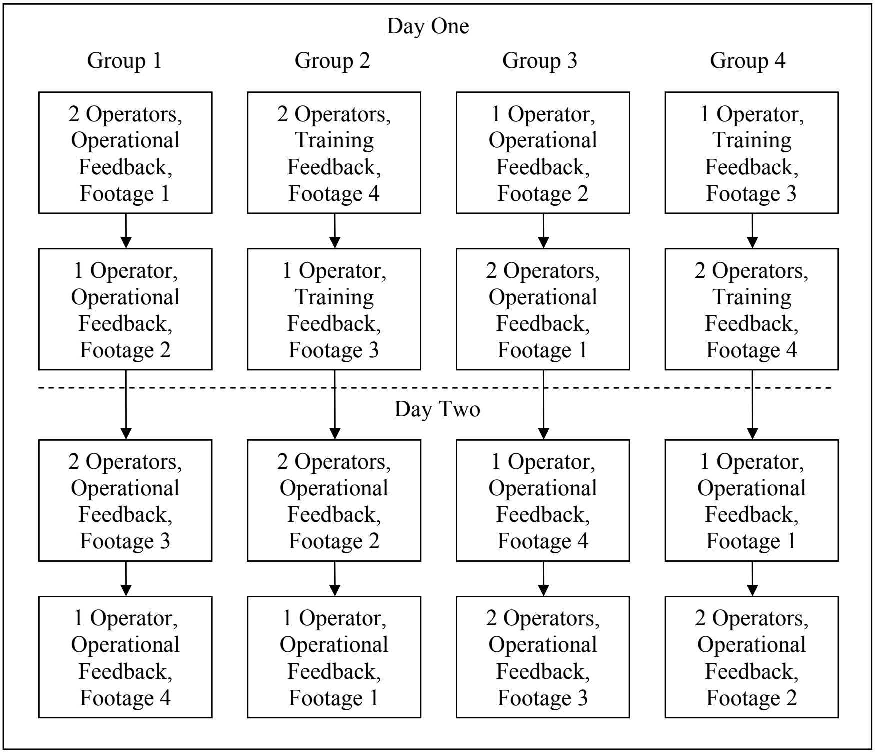


Fig. 5. A graphical outline of the experimental design

were six participants in each group. As indicated in Figure 5, participants completed the experiment both individually, and in a two person group. The ordering of the tasks was counterbalanced, so that some participants completed it on their own first, and some completed it in the group first.

Participants could receive one of two types of feedback. Training feedback involved immediate information regarding the participant's response (i.e., whether their response was correct or incorrect) and informed the participant when a target was missed. Operational feedback involved delayed information on whether or not an identified target was 'carrying', and did not include information regarding missed 'carriers'. The order of type of feedback provided was counterbalanced on the first day, so that half of the participants received operational feedback, and the other half received training feedback. However, on the second day all participants received only operational feedback. This design allowed a comparison of training techniques (training feedback vs. operational feedback) on a final operational test where feedback was similar to that provided in an operational setting.

Before viewing the THz footage, participants were familiarised with the software and task. THz camera still images were shown to participants, comparing ‘carrying’ and non-carrying individuals, as well as some short pieces of footage from the previous trial in which participants were familiarised with the ‘look’ of the software while in operation. This was followed by a practice trial.

Within the experiment proper, participants viewed the footage and were required to make decisions regarding whether the observed ‘passengers’ were carrying concealed drugs or contraband on their person. When participants suspected a ‘passenger’, they were asked to relay the identification number displayed on the chest of the target to the experimenter, along with the suspected location of the item.

3 Results

3.1 Signal Detection Theory Measures

The two basic measures of performance on the THz imaging task are the Hit Rate (HR) and the False Alarm Rate (FAR). The HR is the proportion of passengers who were correctly identified as carrying contraband out of the total number of passengers who were carrying contraband. The FAR is the number of passengers who were incorrectly identified as carrying contraband out of the total number of passengers who were not carrying contraband. The HR and FAR can also be used to calculate two additional measures. The Missed Detection Rate (MDR) is the proportion of passengers carrying contraband that were not identified and is equivalent to (1 – HR). Similarly, the True Reject Rate (TRR) is the proportion of passengers not carrying contraband that were not identified. The FAR is equivalent to (1 – TRR).

In this study, we will also use two other measures derived from Signal Detection Theory (Green & Swets, 1966) to characterise human performance. The first of these measures is known as *discrimination* and relates to a person’s ability to distinguish images of people carrying contraband from those of people who are not. Such an ability may be strongly linked to basic visual ability on the task. The second is known as *bias* and relates to a person’s tendency to respond ‘yes’ or ‘no’ on the task and can be considered a type of decision threshold. For example, a person may have a strong tendency to raise an alarm and will do so even when they do not have a great deal of visual proof that contraband is present. Variables may affect bias and discrimination separately. This may include personality and task effects, e.g., if intelligence has been received that a person may attempt to smuggle contraband on a particular flight then an operator may err on the side of making too many false alarms in the hope of identifying the suspect.

In this report we will use A' and B'' as measures of discrimination and bias respectively (Stanislaw & Todorov, 1999). These particular measures were chosen because they are non-parametric, i.e., in theory, they do not make assumptions about the underlying response distributions (however for a more detailed comparison of the assumptions behind purportedly parametric and non-parametric discrimination and bias measures see Pastore, Crawley, Berens & Skelly, 2003). Previous research has shown that such assumptions in both machine and human matching performance are often incorrect (see Butavicius, 2006 and Fletcher, Butavicius & Lee, 2008).

The values of A' and B'' are calculable directly from the conventional error rates of HR and FAR:

$$A' = .5 + \left[sign(HR - FAR)\frac{(HR - FAR)^2 + |HR - FAR|}{4\max(HR, FAR) - 4(HR)(FAR)} \right]$$

$$B'' = sign(HR - FAR)\frac{HR(1 - HR) - FAR(1 - FAR)}{HR(1 - HR) + FAR(1 - FAR)}$$

An A' value of 0.5 means the images of people carrying contraband cannot be distinguished from people not carrying contraband. A value of 1 equates to perfect performance and values less than .5 possibly reflect sampling error or response confusion. An example of response confusion is where a screener consistently responds that people are carrying contraband when they in fact are not and not identifying contraband when in fact they are. This may suggest that they actually can distinguish between contraband / no-contraband cases but have misunderstood the instructions of the experiment. The minimum possible value of A' is 0. B'' values range from -1 (extreme bias towards raising an alarm) to 1 (extreme bias towards not raising an alarm). A B'' value of zero indicates no response bias.

3.2 Observations on *B''* and Conversion to |*B''*|

In this study, the values of B'' were highly biased towards positive values with only five participants producing negative values. This small number of negative values did not permit separate analyses for positive and negative bias. Therefore, to allow a more parsimonious analysis, the B'' value was converted to a magnitude only. According to this transformation, low values of $|B''|$ reflect less response bias while larger values reflect greater response bias (regardless of whether this is towards 'yes' or 'no' responses).

3.3 REI Scores

We compared the scores we obtained on the rationality and experientiality inventory (REI) to those in the large sample (N = 398) study of Pacini and Epstein (1999). In our study we observed, on average, higher Rationality (4.11 [SD = .46] vs. 3.39 [SD = .061]) and lower Experientiality (3.06 [SD = .56] vs. 3.52 [SD = .042]). This bias may reflect the nature of the scientists who formed the majority of the population sample in our study. REI norms for professional screeners in airport scenarios were not available.

3.4 The Influence of One versus Two Operator Conditions

All participants completed the experiment in both individual and group conditions. Performance scores were examined to evaluate whether detection ability differed between the two conditions. Since the sample size was small (N = 24) and the distributions were not normal, the Wilcoxon Signed Rank Test was used. The performance measures obtained by the group were compared to the average of the individual scores obtained by both group members. Participants completed the task in a different group on the second day of the experiment, and therefore, the results for both days were examined separately.

An examination of the raw data found a slight advantage when participants completed the task in a group compared to when it was completed individually. Groups tended to obtain higher hit rates, higher A' scores, fewer false alarms and lower $|B''|$ scores. When tested using a two-tailed Wilcoxon Signed Rank Test, the differences were all non-significant. However, examination of the effect sizes, as shown in Table 1, show some evidence of meaningful trends in the data. An effect size is independent of statistical significance and its square estimates the variation accounted for by this manipulation as a proportion of the total variation observed. For example, the reduction in response bias on the first day of testing for cases where people worked in pairs as opposed to when they worked by themselves accounted for 12% of the total variation. This figure is not insubstantial given the multitude of other factors influencing performance on the task.

3.5 The Influence of Feedback

Since it is possible that effective feedback could improve future detection performance, the experiment consisted of two types of feedback. On the first day of testing, half of the participants received training feedback. The participants assigned to this treatment condition will be referred to as the ‘Training’ cohort. The other half of the participants received operational feedback. These participants will be referred to as the ‘Operational’ cohort. In order to determine whether participants were assisted by the training feedback, on the second day of testing, all participants received their feedback in an operational style.

In order to allow a valid analysis of any differences in participants’ performance across the two days of testing, the participants viewed the same footage on the first and second day. This was important, since any differences in the difficulty of the footage blocks could create a confound, and jeopardise the ability to ascribe differences in performance to the type of feedback received.

The mean differences in participants’ performance between the two days were examined. The raw scores, contained in Table 2, indicated an advantage for the participants who had received training feedback on the first day. For example, participants who received training feedback on the first day of testing had lower false alarms rates and higher hit rates on the second day. In contrast, the participants who received operational feedback on both days had more false alarms rates and lower hit rates on the second day. Similarly, discrimination (A’) improved across the two days for those who received training feedback and who operated as individuals but there was no evidence of improvement for those who received only operational feedback.

Table 1. Descriptive statistics and Wilcoxon Signed Rank Test results for Individual and Group Performance

<table>
<tr><th></th><th>1st Day of Testing
Mean diff (Individual – Group)</th><th>SD</th><th>Z</th><th>Sig.</th><th>Effect size (r)</th></tr>
<tr><td>HR</td><td>-3.99</td><td>33.61</td><td>-0.43</td><td>.67</td><td>-0.13</td></tr>
<tr><td>FAR</td><td>2.67</td><td>7.65</td><td>-1.02</td><td>.31</td><td>-0.29</td></tr>
<tr><td>A’</td><td>-.05</td><td>.17</td><td>-0.86</td><td>.39</td><td>-0.25</td></tr>
<tr><td>|B’’|</td><td>.20</td><td>.57</td><td>-1.18</td><td>.24</td><td>-0.34</td></tr>
<tr><td></td><td>2nd Day of Testing
Mean diff (Individual. – Group)</td><td>SD</td><td>Z</td><td>Sig.</td><td>Effect size (r)</td></tr>
<tr><td>HR</td><td>-2.26</td><td>30.68</td><td>-0.18</td><td>.86</td><td>-0.05</td></tr>
<tr><td>FAR</td><td>2.95</td><td>7.70</td><td>-1.17</td><td>.24</td><td>-0.34</td></tr>
<tr><td>A’</td><td>-.03</td><td>.15</td><td>-0.54</td><td>.58</td><td>-0.16</td></tr>
<tr><td>|B’’|</td><td>.13</td><td>.53</td><td>-0.16</td><td>.88</td><td>-0.05</td></tr>
</table>

Despite these mean differences in error rates and discrimination there was substantial variation in the individual scores as evidenced by the relatively large standard deviations. The perceived advantage of training feedback was assessed using a two-tailed Wilcoxon Signed Rank Test, and although the differences were all non-significant, for the participants operating by themselves and who received training feedback, the mean difference in the hit rate across the two days ($Z = -1.9$, $p = .053$, $r = -.39$) and *A′* score ($Z = -1.88$, $p = .060$, $r = -.38$) both approached significance, with improved performance on the second day of testing. The increase in performance attributed to the difference in feedback accounted for 15% of the variation observed.

Table 2. Mean differences (Day 2 – Day 1) in performance with standard deviations shown in brackets

<table>
<tr><th></th><th>Condition</th><th>FAR</th><th>HR</th><th>A’</th><th>|B’’|</th></tr>
<tr><td>Individual</td><td>Training</td><td>-1.95 (8.73)</td><td>9.38 (15.19)</td><td>.08 (.13)</td><td>0 (.28)</td></tr>
<tr><td></td><td>Operational</td><td>1.45 (5.93)</td><td>-5.9 (19.66)</td><td>.03 (.08)</td><td>.08 (.28)</td></tr>
<tr><td>Group</td><td>Training</td><td>-4.01 (7.52)</td><td>2.08 (21.69)</td><td>.02 (.06)</td><td>.01 (.38)</td></tr>
<tr><td></td><td>Operational</td><td>1.66 (11.93)</td><td>-4.17 (12.91)</td><td>.02 (.13)</td><td>.05 (.28)</td></tr>
</table>

However, on account of the great deal of individual variation in performance, and given the small sample sizes, it is necessary to interpret these results with caution. Furthermore, there are large differences in the mean scores with far higher means for individual performance with operational feedback and group performance with training feedback, and far lower mean scores for individual performance with training feedback and operational group performance. This is associated with the differences in the footage examined, since participants found it easier to discriminate when viewing certain blocks of footage. Hence, it was more appropriate to assess the mean differences, which provide a valid comparison measure.

3.6 The Influence of Personality on Performance

Personality measures were correlated with performance measures to determine if there were any significant relationships. According to the theoretical background detailed in the introduction to this chapter, individuals who showed a strong inclination towards either rationality or experientiality were hypothesised to have inferior performance on the THz imaging task than those who were not (i.e., were equal on measures of rationality and experientiality). To capture this, a new score, $|R\text{-}E|$, was calculated which represented the magnitude of the difference between an individual's rationality and experientiality scores (i.e., the absolute value of the difference between the two scores). Individuals with a strong tendency towards one characteristic produced larger scores (maximum = 4) than those who did not (minimum = 0). In this chapter, the description of effect sizes as 'small', 'medium' and 'large' for all the personality and performance measures follows the conventions described in Cohen (1988).

Although the results were not statistically significant they did follow an expected trend, but the magnitude of these effects was small. For individuals with a strong preference for one personality characteristic (i.e., a high $|R\text{-}E|$ score), there was a reduction in discrimination ($r(24)= -.1$, $CI_{95\%} = [-0.48, 0.32]$, $p=.64$) and an increase in bias ($r(24)= .29$, $CI_{95\%} = [-0.13, 0.62]$, $p=.17$).

3.7 The Influence of Visual Abilities on Performance

As described previously, participants completed three perceptual tasks; a test of the global precedence effect (Miller, 1981), a feature analysis test and a signal detection task based on Landolt's (1888) broken ring task. It was hypothesised that the participants who obtained higher scores on those tasks would have better detection performance.

The relationship between perceptual ability (as measured by the global precedence, feature analysis and signal detection tasks) and THz task performance (as measured by hit rate, false alarm rate, A' and $|B''|$ scores), were measured using a series of Pearson product moment correlations. There was a medium positive correlation between participants' average performance on the global precedence task and their hit rate ($r\ (19) = .5$, $CI_{95\%} = [0.06, 0.78]$, $p < .05$) and A' score ($r\ (19) = .54$, $CI_{95\%} = [0.16, 0.82]$, $p < .05$) in the THz task. Furthermore, there was also a medium positive correlation between the time taken to complete the global precedence task and the percentage of false positives in the detection task ($r\ (19) = .51$, $CI_{95\%} = [0.07, 0.79]$, $p < .05$). Performance on the global precedence task can be further broken down into the global and local subtasks, i.e., where the participant is required to attend to global features when the local features are distractors, and vice versa. There was a significant positive correlation between discrimination ability on the THz task and the global task ($r\ (19) = .57$, $CI_{95\%} = [0.16, 0.81]$, $p < .05$) and the correlation with performance on the local task was similar in size and close to statistical significance ($r\ (19) = .37$, $CI_{95\%} = [-0.01, 0.75]$, $p = .06$).

However, no such relationship was evident between performance on the THz task and the other tests of visual perception.

4 Discussion and Conclusions

4.1 One versus Two Operator Conditions

Participants completed the experiment in both individual and group conditions, with the aim of determining whether either condition provided a performance advantage. Although the observed differences were not significant, the results indicated an advantage in the group condition that was evident in both an increase in discrimination and a reduction in response bias. Previous literature suggests that group decision-making often results in group polarisation, in which decisions made by a group are more extreme than decisions made by individuals (Wallach et al., 1962). Based on this theory, it is possible that, when there was uncertainty surrounding the presence of contraband, groups may have been more likely to generate an alarm (or withhold an alarm response) than individuals.

It is, however, necessary to note that there was large variation in group performance. For example, observations from the experiment indicated that some groups were more likely to be distracted. In fact, there were situations when groups missed targets because they were not focused entirely on the passengers. In contrast, there were other occasions when the groups hardly spoke, and there were also occasions when one participant took a more dominant role, and tended to make the decisions. There were also some examples of deductive reasoning between participants, where hypotheses (such as the theory that a targets' belt was obstructing the view of potential contraband) were discussed between participants to either verify or disprove thoughts.

Some of these differences could be influenced by the relationship between the participants. For instance, the interaction is likely to be different when a supervisor and subordinate are grouped rather than when two colleagues of equal status are grouped. Unfortunately, in this experiment the number of pairs was insufficient to make reliable conclusions regarding the influence of these factors.

4.2 The Influence of Feedback

A further aim of the study was to examine the performance differences that arise from operational feedback versus training feedback. The participants who received training feedback on the first day tended to have improved performance on the second day, whereas the participants who received operational feedback on the first day tended to have worse performance on the second day. This pattern was strongest when participants were working individually. This suggests that the information regarding missed detections was very useful, and improved future performance. Theoretically, the feedback scheme better allows an operator to fine-tune their perceptual processing than on-the-job training. However, again these findings were not significant, and further investigation with larger sample sizes is necessary to determine whether this finding is reliable.

4.3 The Influence of Personality Factors

It was hypothesised that individual differences in personality may lead to differences in performance outcomes. As outlined in section 2.3.3, Cognitive-Experiential Self-Theory emphasises that information processing is dependent on personality and can be broadly classified as either rational or experiential. Individuals who favour a rational approach analyse information in a more conscious and deliberate manner, whereas experiential information processing tends to be characterised as more automatic, spontaneous and natural (Epstein, 1994).

Research also suggests that individuals who are experiential are more impulsive than individuals who are 'rational', and due to higher levels of impulsivity, 'experiential' individuals are also more error prone. Although rational individuals are more likely to have a lower error rate, as a result of lower impulsivity, they tend to miss a lot of information because of the extended time it takes them to process that information (Dickman, 1985). In the THz task, when such individuals are deliberating on whether a passenger is carrying contraband or not, they may be more likely to miss contraband on passengers who subsequently pass by the camera.

The observed relationships between personality measures and performance in this study were not statistically significant. In addition, the size of the trends observed was less than that found when we examined the link between performance on the THz imaging task and visual abilities. It should be noted that the trends observed were consistent with the predictions made based on the background literature and this has implications for future research. For example, the participants in this study who were high on rationality had fewer hits, poorer discrimination and greater bias. This can be explained by the expectation that rational individuals have lower impulsivity with the result that they are more cautious in their decision making processes. On the other hand experiential individuals were more likely to have more hits, greater discrimination and less bias.

4.4 The Influence of Perceptual Factors

The experiment also found that participants with better performance on the global precedence task (Miller, 1981) tended to have better contraband detection performance. More specifically, those participants who were better able to selectively focus on either global or local features of a visual image performed better on the THz task. It may be that the contraband "shadow" in the image constitutes a localised feature of the entire body image (which may be considered global). Regardless, the general finding suggests that visual processing abilities play an important role in how well an operator can use the THz device to identify contraband. This also has important implications on how to select or recruit individuals as operators on this device. The relationship between basic visual abilities and operator performance on screening devices is worthy of further investigation.

4.5 Limitations

Although this study has provided a very important insight into the human factors issues associated with THz imager performance, there are a number of possible limitations which should be taken into account. Firstly, participants were undertaking an experiment, and the importance of correct detection in an experimental condition is not necessarily akin to the importance of perfect detection in a real-life situation. For example, in the experiment, there was no penalty associated with an incorrect decision, as a false detection did not result in the questioning of an innocent passenger. Similarly, a missed target did not result in actual contraband passing detection and the potential harm or damage that could result. In contrast, in an operational environment, people may be less likely to risk making an incorrect detection. Hence, it is possible that the performance found in this experiment may not directly translate into expected performance in an operational environment.

There were also limitations associated with the footage used. For instance, there were some instances where the ‘passenger’ did not walk directly through the camera, which may have influenced the participants’ ability to make an accurate decision. Furthermore, as there was only a limited amount of footage available, it was necessary for participants to view the same blocks on more than one occasion. This introduces the possibility that participants may have attempted to remember the correct identification numbers. However, since there was always a day in between participants reviewing the same segment, the influence of memory should be minimised.

Finally, it is also important to note that there was a great deal of individual variation, which makes it difficult to draw concrete conclusions. Together with the relatively small sample size, this means that many of the results are reported as trends, and it would be necessary to repeat the experiment with more participants to obtain statistical significance and ensure reliable results.

4.6 Future Directions

This experiment has raised a number of other factors that could be studied in more depth in the future. As mentioned above, individuals varied greatly in their performance, and it would be interesting to further examine the factors that may underly these differences, to better understand the variation in individual performance. This study found an influence of some personality and perceptual factors. Future studies could examine whether individual differences in areas such as attention and memory, cognitive aptitude, expectations and other abilities may influence detection performance. Perhaps the most important suggestion for future experimentation is a greater sample size given the individual variation in performance observed in the current study.

A future experiment could also examine the influence of fatigue on performance. Studies have shown that fatigue (specifically task-induced fatigue) can reduce perceptual sensitivity, impair efficiency, and can reduce individuals’ awareness of their impairment (Brown, 1994; Matthews & Desmond, 2002). Since operators would often be required to maintain attention for long shifts, it is possible that their monitoring performance may be influenced by fatigue. Due to restraints in the amount of footage available, this factor was not assessed in the current study, but future experiments could assess whether performance reduces over long periods of time.

In addition, given the significant contribution that visual abilities play in performance on the THz task, future experiments could examine this relationship in further detail. In particular, a future study involving eye-tracking could help determine the visual properties of the image which a successful operator focuses on to find contraband (see Fletcher, Butavicius & Lee, 2008). This could have a significant impact on training purposes by teaching potential operators the best visual strategies for detecting contraband.

In more general terms, the current study can be used as a template for human factors research in a wider range of airport security devices. This type of research is particularly important where any such device relies on the operator's visual and cognitive abilities for an alarm to be raised. Such research does not fall into the conventional style of evaluation for airport security devices as described in Butavicius (2006). However, as demonstrated in this chapter, such research can address important issues on selecting personnel for such roles and policies on procedures for operating these devices in the airport environment. In turn, such studies can improve the efficiency of the deployment and use of cutting edge screening devices in a world under increased threat of airport related terrorism.

References

Atthowe, J.: Interpersonal decision making: The resolution of a dyadic conflict. Journal of Abnormal and Social Psychology 62(1), 114–119 (1961)

Biederman, I.: Recognition-by-components: A theory of human image understanding. Psychological Review 94, 115–147 (1987)

Brown, I.: Driver Fatigue. Human Factors 36(2), 298–314 (1994)

Butavicius, M.A.: Evaluating and predicting the performance of an identification face recognition system in an operational setting. Australian Society for Operations Research Bulletin 25, 2–13 (2006)

Butavicius, M., Foster, R., Whittenbury, A., Parsons, K., McCormac, A., MacLeod, V.: Human Factors Study of the ThruVision T400 Passive Terahertz Camera. DSTO Technical Note, DSTO-TN-0821 (2008)

Butavicius, M.A., Mount, C., MacLeod, V., Vast, R., Graves, I., Sunde, J.: An experiment on human face recognition for access control. In: Lovrek, I., Howlett, R.J., Jain, L.C. (eds.) KES 2008, Part I. LNCS (LNAI), vol. 5177, pp. 141–148. Springer, Heidelberg (2008)

Cohen, J.: Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum, New York (1988)

Dickman, S.: Impulsivity and perception: Individual differences in the processing of the local and global dimensions of stimuli. Journal of Personality and Social Psychology 48(1), 133–149 (1985)

Endsley, M.R., Jones, W.M.: Situation awareness, information dominance & information warfare. AL/CF-TR-1997-0156, United States Air Force Armstrong Laboratory, Wright-Patterson Air Force Base, OH (1997)

Epstein, S.: Integration of the cognitive and the psychodynamic unconscious. American Psychologist 4, 709–724 (1994)

Fletcher, K.I., Butavicius, M.A., Lee, M.D.: The effects of external features and time pressure on unfamiliar face matching. British Journal of Psychology 99, 379–394 (2008)

Gerrig, R.J., Zimbardo, P.G.: Psychology and Life, 17th edn. Allyn & Bacon, Boston (2005)

Goldstein, E.B.: Sensation and Perception, 6th edn. Wadsworth, California (2002)

Green, D.M., Swets, J.A.: Signal Detection Theory and Psychophysics. Wiley, New York (1966)

Hall, B., Bird, R., Johnson, R., Resnyansky, L., Sunde, J.: Evaluation of the ThruVision T4000 Passive Terahertz Camera in an Operational Environment. DSTO Client Report, DSTO-CR-2007-0288 (2007)

Hardmeier, D., Hofer, F., Schwaninger, A.: The X-Ray Object Recognition Test – A reliable and valid instrument for measuring visual abilities needed in x-ray screening. IEEE ICCST Proceedings 39, 189–192 (2005)

Humphrey, G.K., Jolicoeur, P.: An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification. Quarterly Journal of Experimental Psychology 46A, 137–159 (1993)

Landolt, E.: Methode optometrique simple. Bull. Mem. Soc. Fran. Opthamal 6, 213–214 (1888)

Laughlin, P.R., Bitz, D.S.: Individual Versus Dyadic Performance on a Disjunctive Task as a Function of Initial Ability Level. Journal of Personality and Social Psychology 31(3), 487–496 (1975)

Lee, M.D., Vast, R.L., Butavicius, M.A.: Face matching under time pressure and task demands. In: Sun, R., Miyake, N. (eds.) Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada, pp. 1675–1680 (2006)

Mayer, J.D., Salovey, P., Caruso, D.R.: Models of emotional intelligence. In: Sternberg, R.J. (ed.) Handbook of Human Intelligence, 2nd edn., pp. 396–420. Cambridge, New York (2000)

Marr, D.: Vision: A computational investigation into the human representation and processing of visual information. W.H. Freeman, San Francisco (1982)

Matthews, G., Desmond, P.A.: Task-induced fatigue states and simulated driving performance. The Quarterly Journal of Experimental Psychology 55(2), 659–686 (2002)

Miller, J.: Global precedence in attention and decision. Journal of Experimental Psychology: Human Perception and Performance 7(6), 1161–1174 (1981)

Myers, D.G., Lamm, H.: The group polarization phenomenon. Psychological Bulletin 83, 602–627 (1976)

Navon, D.: Forest Before Trees: The precedence of global features in visual perception. Cognitive Psychology, 353–383 (1977)

McCarley, J.S., Kramer, A.F., Wickens, C.D., Vidoni, E.D., Boot, W.R.: Visual Skills in Airport-Security Screening. Psychological Sciences 15, 302–306 (2004)

Orasanu, J.: Evaluating team situation awareness through communication. In: Garland, D.I., Endsley, M.R. (eds.) Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness, pp. 283–288. Embry-Riddle Aeronautical University Press, Daytona Beach (1995)

Pacini, R., Epstein, S.: The relation of rational and experiential information processing styles to personality, basic beliefs and the ratio-bias phenomenon. Journal of Personality and Social Psychology 76(6), 972–987 (1999)

Palmer, J., McLean, J.: Imperfect, unlimited capacity, parallel search yields large set-size effects. Paper presented at the Society of Mathematical Psychology, Irvine, CA (1995)

Pastore, R.E., Crawley, E.J., Berens, M.S., Skelly, M.A.: “Nonparametric” A’ and other modern misconceptions about signal detection theory. Psychonomic Bulletin and Review 10(3), 556–569 (2003)

Schwaninger, A.: Computer based training: a powerful tool to the enhancement of human factors. Aviation Security International, 31–36 (February 2004)

Schwaninger, A., Hardmeier, D., Hofer, F.: Aviation security screeners visual abilities and visual knowledge measurement. IEEE Aerospace and Electronic Systems 20, 29–35 (2005)

Stanislaw, H., Todorov, N.: Calculation of signal detection theory measures. Behavior Research Methods, Instruments & Computers 31(1), 137–149 (1999)

Treisman, A.: The perception of features and objects. In: Baddeley, A., Weiskrantz, L. (eds.) Attention: Selection, Awareness, and Control, pp. 5–34. Clarendon Press, Oxford (1993)

Vickers, D., Butavicius, M.A., Lee, M.D., Medvedev, A.: Human performance on visually presented traveling salesman problems. Psychological Research 65, 34–45 (2001)

Wallach, M.A., Kogan, N., Bem, D.J.: Group influence on individual risk taking. Journal of Abnormal and Social Psychology 65(2), 75–86 (1962)

Chapter 7
Biorhythmic Analysis to Prevent Aviation Accidents

R.K. Saket[1,*], Wg. Cdr. S.P. Kaushik[2], and Col. Gurmit Singh[3]

[1] Department of Electrical Engineering,
Institute of Technology,
Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India
`rksaket.eee@itbhu.ac.in, saketrk@yahoo.com,`
`drrksaket@gmail.com`
`http://www.itbhu.ac.in, http://www.bhu.ac.in`
[2] Department of Mechanical Engineering,
Sobhasaria Engineering College, Sikar (Rajasthan), India
`spkaushik@yahoo.co.uk`
[3] Faculty of Engineering and Technology,
Sam Higginbottom Institute of
Agriculture, Technology and Sciences–Deemed University,
Allahabad (Uttar Pradesh), India
`gurmitsingh3@rediffmail.com`

Abstract. The Word "*Biorhythm*" is derived from the Greek word "*Bios*" which means life and "*Rhythm*" which means flowing with a regular movement. Biorhythm theory uses mostly scientific methods to chart the rhythms or cycles that affect the internal functioning of the body and of human behavior. This is particularly applicable to the physical, emotional and intellectual or mental abilities. Biorhythm theory states that, at the moment of birth three statistical cycles are initiated and these will recur consistently throughout life. This chapter proposes an investigation of the reasons for human error as a contributing factor in flying accidents. Physical factors such as man machine interactions involved in flying mistakes and a taxonomic approach to errors has been proposed in this chapter, in order to avoid accidents. This chapter presents a new methodology based on a probabilistic approach for biorhythmic analysis in an attempt to prevent aviation accidents. The methodology has been developed using a Gaussian distribution technique for evaluation of the aviation system reliability considering the biorhythmic effects on the pilot. Normal distributed data from the US air force were tested and analyzed. These were based on the performance ability of pilots and the peak demand of the performance using a Gaussian distribution approach. Validation of an aircraft accident due to biorhythm is explained in this chapter with consideration of the peak performance demand and differing standard deviations in the performance ability of each pilot. A new curve named the Incident – Duration Curve has been introduced. This is based on a biorhythmic analysis of Indian and US air force data. The area under normal distribution curve of the US air force data represents the successful performance ability zone

---

[*] Corresponding author.

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 207–240.
springerlink.com

of the Pilot. The accident zone is the area of operation during which the Performance Demand exceeds the Performance Ability of the particular aircraft pilot. Operation within the zone of the normal distribution curve is successful owing to ability and fitness of the particular pilot. Failure probabilities considering Peak Performance Demand and pilot's ability have been evaluated using a Gaussian distribution approach. A Safety Factor Concept is also given in this chapter. This has been done so as to include biorhythmic analysis in the attempt to avoid aviation accidents. A Stepped Incident-Duration Curve has been utilized in order to evaluate the particular pilot's reliability when using the system. The complete aviation system was evaluated by using Simpson's 1/3rd rule.

Keywords: Biorhythmic Analysis, Incident-Duration Curve, Performance Ability of that pilot included items such as: Performance Demand, Accident Prone Zone/Critical day, Gaussian distribution, Reliability Evaluation, Safety Factor.

1 Introduction

The word *Biorhythm* is derived from the Greek word *Bios* means life and *Rhythm* meaning a regular moment. Biorhythm theory mostly uses scientific methods to chart the rhythms or cycles that affect the internal functioning of the body and human behavior in particular based on the physical, emotional and intellectual or mental abilities. The biorhythm theory states that from the moment of birth three statistical cycles are initiated and they recur consistently throughout a person's life. Our Productivity, Efficiency, Intelligence and activity levels are not merely just matters of will power. Thommen George S. (1973) has explained that each of us is subjected to biological rhythms. The three biorhythm cycles have independent duration and influence. These cycles compose the classical theory, which became popular with the general public after 1970. The earliest observed biological cycles were recorded by Alexander the great scribe, Androsthenes, in the fourth century BC. Many studies done abroad, including the United States during the 1940's and 1950's have demonstrated a higher disposition towards accidents and human error that coincide with these biorhythmic cycles. The physical, emotional and intellectual biorhythm cycles have sinusoidal characteristic and are shown in fig. 1. They are of 23, 28 and 33 days duration, respectively. At birth these cycles begin at zero and then follow the above sinusoidal characteristics. This characteristic goes in the plus direction and return to zero or mid cycle, then go in the minus or downward direction. They then turn around and return to the positive. The cycle is then repeated. Biorhythm cycles are thus composed of positive phases, negative phases and nodal points where the curve crosses the abscissa. Hines, Terence M. (1998) state each cycle starts on the positive phase at the moment of birth. The positive phases correspond to periods of better performance and the negative phase corresponds to periods of poor performance and the greatest susceptibility to harm. Critical periods are usually 24 hrs to 48 hrs duration. The physical cycle has a 23 days period and affects a broad range of physical factors like resistance to disease, strength, coordination, speed and other basic body functions. In addition to a sensation of physical well being, the emotional cycle has a period of 28 days and affects creativity, sensitivity, mental health and mood for example. The intellectual cycle has a period of 33 days and it affects memory, alertness, receptivity to knowledge and logical analytical functions of the mind. Since the periods of all three cycles are different (23, 28 and 34 days), the interaction of the three cycles when overlaid on each other is complex. They are not in this exact configuration for a further 21, 252 days or 58 years and 66-68 days, depending on leap years. The theory of Davis W. Carvey and Roger G. Nibler (2006) predicts that accidents will occur more on accident prone days. On these days more than one cycle, out of the three will cross the abscissa. Zimmerman, R. O. (2001) has presented an example of a study related to biorhythm. This was popular in the late 1970's it is used to illustrate the separating of scientific evidence and pseudo-science. The cited biorhythm study focuses on the relationship of the accident dates and the three biorhythm cycles. John C. Aldrin; Entique A. Medina; Daniel A. Allwine; Mohammed Qadeer Ahmed; Joseph Fisher; Jeremy S. Knopp; and Eric A. Lindgren (2006) have described human factors in a nondestructive evaluation which is critical to maintain inspection reliability. Reliability of structural health monitoring systems is particularly sensitive to sensor degradation over time. To investigate the impact of these issues, probabilistic models for risk assessment and cost benefits analysis have been developed. Quantitative studies are presented evaluating the effects of variations in probability of detection associated with human factors (Michael Clarke, Editor: 1995). They also evaluate the in situ sensor degradation of life cycle measures which include factor of cost and the probability of failure.

This chapter is organized to take account of pilot error in aircraft accidents, the human factor component in aircraft accident analysis, the human factor in accident patterns, statistical analysis of the accident data and Reliability Evaluation of the system based on different probabilistic analysis. In Section: I the meaning of the word 'Biorhythm' and its brief history is given. Section: II gives an overview of typical natural biorhythm cycles and its considerations - physical, emotional and intellectual cycles in detail. The causes of aircraft accidents including direct and indirect causes are highlighted in section: III. Behavioral analysis and Impact of cycles of biorhythm on human performance with causes and safety measures are illustrated in section: IV. Analysis of causative Human Error Factor and a Developmental Model for Predicting Human Error where consideration to various causes have been described in section V. Statistical techniques for biorhythmic analysis for accident prone days are described in Section: VI. In Section: VII the reliability analysis of the biorhythmic aviation accidents is discussed. Several methods of analysis are given in this section. They include Gaussian Distribution Approach, Failure Probability Evaluation Approach, Safety Factor Concept and Peak Load Considerations Approach and Performance Evaluation Approach using Simpson's 1/3rd rule are presented. Section: VIII represents results and discussions based on biorhythmic data. Conclusions and projected future research work plans are given in section: IX.

2 Typical Biorhythm Cycles: An Overview

As previously stated at the moment of birth each of the three biorhythm cycles is initiated and the cycles proceed to follow a fixed sinusoidal pattern throughout the life of the individual. The physical cycle P has a period of 23 days, the sensitivity (or emotional) cycle S has a period of 28 days and the intellectual cycle has a period of 33 days. Individuals with different birth dates will consequently have different composite biorhythm charts, although the theory holds that the cycles of all individuals follow the same 23, 28 and 33 days natural biological rhythm. The calculation of an individual's biorhythm at any given time requires that the date being investigated is specified. The subject's age in days from the date of birth, up to an including the date of interest must next be determined [England, C.E. and Naitoh, P. (1980)]. In this calculation individual biorhythm consideration need to be given in terms of regular leap years and centurial leap years.

The equations for the natural biorhythm curves are as follows:

(a) Physical Biorhythm Cycle: $\sin(2\pi t / 23)$,
(b) Emotional Biorhythm Cycle: $\sin(2\pi t / 28)$,
(c) Intellectual Biorhythm Cycle: $\sin(2\pi t / 33)$,
(d) Intuitive Biorhythm Cycle: $\sin(2\pi t / 38)$

Where, t indicates the number of days since birth. Typical natural biorhythm cycles are shown in figure (1). Percentage values indicated in Y-axis of fig (1) represent percentage numerical values of physical, emotional, intellectual and intuitive natural biorhythm cycles.

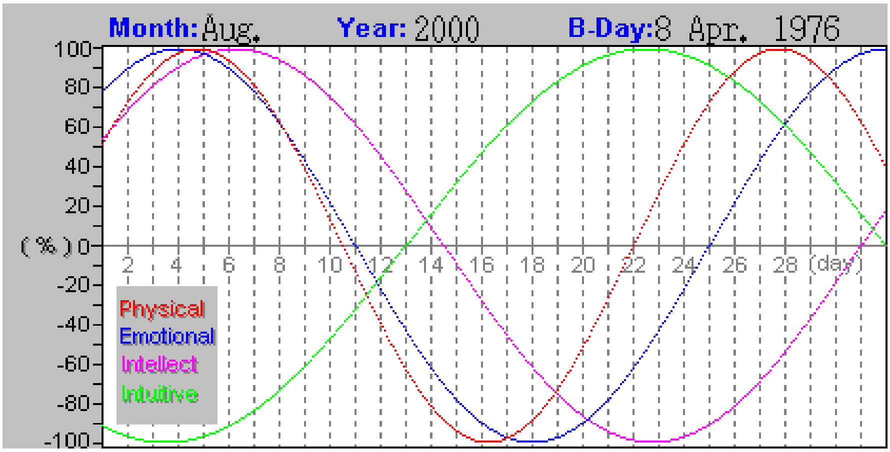


Fig. 1. Natural Biorhythm Cycles

(a) The Physical Biorhythmic Cycle (23 days)

The physical cycles originate in the muscle tissues or fibers. The physical cycle is from our masculine inheritance and affects our physical condition. During the plus side of the cycle (day 2 through day 11) our physical condition is in a charged state. That is we are optimistic, our stamina is high, we need a good deal of movements and physical work is easier. We feel more vigorous and have more vitality. Our endurance level is higher and this is therefore a time of activity and for starting new tasks. Some doctors believe that the days 2 through to 9, that is in the plus half of the cycle are the best days to have elective surgery. During the minus portion of the cycle which is day 13 to day 23 is a recuperative recharging state and one may tire more easily. This period is conducive to recuperation and we are less resistant to stress and physically activity. This is not a good time for starting difficult or energy demanding tasks. George S., Thommen (1973) states that some athletes, depending on this and the state of other cycles and factors can experience a slump during this time. Despite this a well trained athlete who has not over prepared may succeed at this time. It is not a "bad" time. In fact it can be a good time to practice routine physical activities and to "recuperate." Thommen compares the physical cycle to that of a car battery and generator. The fully charged battery can spark the ignition to provide full power. When the battery has run down the generator switches to the charge mode and returns the battery to full strength. The critical points in the physical cycle are day 1 and day 12 1/2. We may be more prone to misjudge our physical energy or endurance while switching from one phase to the other. On critical days we must be more careful, more attentive, and should not hesitate to put off things which involve much physical efforts.

(b) The Emotional Biorhythmic Cycle (28 days)

The emotional cycle governs the nervous system. It is due to the influence on the nerve cells from one's feminine inheritance and it affects the emotional level. During the high end of the cycle (day 2 to day 14) one is more inclined towards optimism and cheerfulness. Creativity, productivity, friendship, feelings, love and cooperation are favorably influenced. The positive phase brings optimism, joy, openness, tolerance, and self control. During the low end of the cycle (day 16 to day 28) your emotions are in a recuperative state as explained by Hines, Terence M. (1998). You are more inclined to be irritable and negative. The negative side brings pessimism, withdrawal, bad moods and sometimes completely illogical sadness. The relative highs and lows of these two phases is definitely influenced by our general temperament. An excitable person will have a wider swing than a more sedate or calm person. The critical days are day 1 and day 15. Insurance and industrial statisticians in the US and in other lands have noticed a higher percentage of self caused accidents on these days. Drivers and other people needing to react quickly with sound judgment should be cautious on these days. According to George S. Thommen, (1973) there is something interesting about these critical days. Since the emotional cycle is 28 days long, exactly 4 weeks, day 1 and day 15 always fall on the day of the week that you were born. Every other week, this day is a critical day in your emotional cycle. If you don't know what day of the week you were born your bio-chart can tell you. Just look at the days when your emotional cycle is on the axis between plus and minus.

(c) The Intellectual Biorhythmic Cycle (33 days)

The Intellectual Cycle was not discovered together with the physical and emotional cycle and it does appear to have less prominence than the other two. It does, however, have an influence. The intellectual cycle originates in the brain cells. When the intellectual cycle is on its high, on its plus phase (day 2 to day 16) we are more capable of absorbing new ideas and appear to think more clearly. Mental responses are more spontaneous and memory functions better according to Douglas, Neil and Francis, L Sink (1976). In the positive phase we have maximum powers of concentration and our memory skills are high. We can adapt to any situation and can make difficult decisions during this period. This is a good time for creative thought and the study of new ideas. During the low phase (day 18 to day 33) your capacity to think may be reduced. This may be a better time to rehearse and review known concepts. Practice of things known will facilitate the storage into the mind and the sub-conscious. The critical points are at day 1 and day 17 1/2. On these days we should defer making important decisions.

2.1 Interpretation of Biorhythm Cycles

Biorhythmic study focuses on physiological, emotional and intellectual processes and their forecasting. Biorhythm phenomena are observable human conditions and can be detailed and explained by biorhythmic studies. Each cycle oscillates between a positive phase [0% to 100%] and a negative phase [- 100% to 0%], during which time bioelectric activity strengthens and weakens. In the waveform of any of the three cycles, the positive period is thought to represent favorable conditions; they are high performance intervals for intellectual function (I) or for physical coordination (P). On the other hand the negative period is thought to represent a recharging phase. During the recharging phase it is believed that a person is inclined to tire more easily (P) become depressed or irritable more readily (S), thus exhibiting a lesser degree or acuity in the learning and decision making process. In the workplace - railways, roads and airlines have most experimented with biorhythm. A pilot describes both the Japanese and American attitude towards biorhythms [Shaffer, J. W.; Zlotowitz, H.L and Fisher, R. S. (1978)]. He acknowledges, researching his pilot logbook, and finding that his greatest errors of judgment occurred on critical days. He concludes that an awareness of one's critical days and the need to pay extra attention to the matter at hand is essential to ensure safety [Khalil, T.M. and Kurucz, C.N. (1977)].

2.2 Accident Prone Zone / Critical Days

Critical days have been described as being full of danger and difficulties. They are accident-Prone days and are said to occur when one's energy expenses change from the positive phase to the negative phase or vice versa. These days also called Critical Days and are considered to be accident-prone because the body's system is in a state of transition and are not stable. They are days of flux and great instability. This instability does not in itself cause accidents but does apparently have a mild negative influence on the performance, which may increase danger. Critical days are not the days when an accident will occur, but are a time when you will be more accident prone. John, H. Wolcott; Mc Meekin R. Robert; Burgin R. E.; and Yanowitch R. E. (1977) have represented the correlation of general aviation accidents with the Biorhythmic Theory and the experience of it as accident-prone days. On these days the organism polarity is in state of flux and therefore the feedback process is highly variable. In this period the organism does not experience immediate and accurate assessment of its capacity. Each accident case was analyzed to determine whether or not the accident occurred on a biorhythmically critical day as shown in table I. The data was also systematically evaluated for the existence of Non- Biorhythmic Cycles. Accidents can be prevented if an individual is prevented from working in a hazardous situation on critical or accident-prone days.

2.3 Combined Biorhythm Cycles

The three-biorhythm cycles may be charted on one curve as shown in Fig. 1. The three rhythms are plotted independently and their relative positions will change from month to month. Combined biorhythm cycles illustrate several accident-prone days. Accident-prone day 1 shows a crossing of the Physical Cycle. Here the sensitivity cycle is recharging, while the intellectual cycle is at a high. On the other hand, an accident-prone day illustrates that a double-crossing of the zero axes appears to be a Double Critical Day. Similarly, three crossings within the same day results in a Triple Critical Day.

3 The Causes of Aircraft Accidents

The correlation of occurrences of aircraft accidents to the critical and negative phases of the biorhythm cycles have been investigated by John, H. Wolcott; Mc Meekin R. Robert; Burgin R. E.; and Yanowitch R. E. (1977). Data from 880 US air force pilots involved in accidents were studied and added to 4278 previously reported cases. The data were tested by Chi-Square analysis under the null hypothesis that proposed there is no effect of biorhythm on aviation accidents. Using this hypothesis, the expected number of accidents occurring on critical days should be 179.13 for the US air force. The investigation of Sacher, Dietmar (1974) dealt critically with the problems of biorhythmic and its influence on human error and accidents. This was based on data obtained from 4346 naval aircraft mishaps in the Fiscal year 1968-1973. John, H. Wolcott; Mc Meekin R. Robert; Burgin R. E.; and Yanowitch R. E. (2006) have calculated biorhythm for over 4000 pilots involved in general aviation accidents in 1972. The data was obtained from the National Transportation Safety Board. The data was analyzed for a correlation of aircraft accident occurrences with both biorhythmically critical days and with individual and multiple low or negative phases of cycle. The causes of aircraft accidents in military aviation can be classified into Direct and Indirect causes.

3.1 Direct Causes of Aircraft Accidents

Direct causes of aviation accidents [Scott A. Shappell; Cristy A. Detwiler; Kali A. Holcomb; Carla A. Hackworth; Albert J. Boquet; Douglas A. Wiegmann (2006)] are shown to be directly responsible for the aircraft accidents. Direct causes are sub-classified as follows:

3.1.1 Technical Defects in Aircraft

Technical defect indicates failure of some aircraft system while it is flying. Example of failure chances for a single engine of twin-engine aircraft system is the failure of the carriage system to come down. Technical defects create hazardous situations, which may lead to an aircraft accident [Kaushik, S.P. Murthy, W. Selva and Shrivastava, K.K. (1990)]. If crew fails to take proper steps to deal with this hazardous situation when landing the aircraft, safety for the nearest aircraft is also impaired. The reliability of the aircraft system where accidents are due to technical defects are evaluated using a Binomial Distribution.

3.1.2 Environment Factors

Environment factors are the factors, which are beyond the control of the pilot/crew/military aviation. Some typical environmental factors are:

- ***3.1.2 (A) Bad Weather Conditions:*** Some military aircraft have to fly in bad weather conditions. For example CB clouds creates hazardous environmental situation. CB clouds are charged clouds which may jam the Gyro instruments in the aircraft. Severe updraft and down draft which are inherently present in the cloud can throw the aircraft, up or down until it disintegrates. The best way to overcome this hazardous situation is to avoid flying into this type of cloud. The exact position of the CB clouds should be given to the pilot by Weather Radar and Ground Control.
- ***3.1.2 (B) Bird Strikes:*** The aircraft needs to avoid bird strikes during takeoff and landing. The aircraft is flying at supersonic speed and birds act as a missile which will damage the aircraft. George E. Meyer (1974) has presented activities related to the aircraft hazard as a result of bird strike at Charleston AFB, South Carolina. He studied a 500 square mile coastal area from 1 June 1971 to 1 June 1972. He gave the theoretical development for the calculation of the Binomial Probability Distribution Functions useful for assessing the risk of bird hazards to aircraft by the use of radar. Each distribution function has been studied in order to determine the degree of risk and the corresponding number of birds involved. The cumulative probability of bird strikes over an entire route can be determined by calculating the union of discrete cell probability sets.

3.1.2 (C) Ricochets**:** Ricochet is an environment factor in military aviation. Ricochets occur when bullets fire from an aircraft are reflected. The reflected bullets may hit the originating aircraft or one of the other aircraft in the formation during bombing or target practice.

3.1.3 Human Factors

Human errors are basic mistakes committed by the Pilot or the Aircrew during flight. It was observed that even the most experienced pilot had committed basic mistakes of landing with the flaps in up position calling three greens when the under carriage lever is in the up position. These human errors can lead to a major aircraft accident if not detected in time. It is difficult to understand, why experienced pilots have made such silly mistakes. Psychologists believe that the mistakes committed by the experienced pilots may be due to some indirect causes. These effect the functioning of the mind of the pilot and interface with his skills. This ultimately leads to the commitment of such mistakes.

3.2 Indirect Causes of Aircraft Accidents

The indirect cause is some factor, which affects the human performance. It can lead to an aircraft accident. The indirect cause is an inter-action between skill and stress. The pilots acquire the skill of flying an aircraft through an intensive training for a minimum of a three year period. Experienced, pilots have capabilities to overcome unfavorable environmental situations and even to deal with technical defects when developed during flying. The performance of these pilots while flying an aircraft is the result of an interaction between skill and stress. Skill represents physical and mental capabilities. Acquired knowledge about the aircraft and its operations and factors such as experience are also important. Stress is a feeling of hardship or tension caused by an over powering situation when the individual feels that his resources to deal with it are inadequate. Stress is a part and parcel of a human being and is unavoidable. Pilots have to routinely face stresses. The pilots are subjected, to two types of stress.

3.2.1 Cumulative Stress

(A) Unusual Life Condition: Unusual life conditions may include an unhappy family life, financial problems, and the frequent transfer of pilots. A study by NATO revealed that cumulative stress loads of such events in the immediate past predispose a person to psychosomatic or a purely physical cause for a reduction of ability (Thompson, Simon G.; Ghanea-Hercock, Robert: 2005). Such cumulative stress may lead to attention failure, error of judgment or forget fullness.

(B) Life Style and Temperament: Life style factors are (a) Over ambition, (b) Constant Worry, (c) Expectation of Perfection in every event. A poor temperament of the pilot is one which is responsible for some air craft accidents.

(C) Zero Error Factor: Modern aviation undoubtedly calls for a zero error factor. This is an inescapable requirement. Some social scientists believe that it is impossible to achieve a zero error factor as a person is likely to make a mistake at some time. Achievement of zero error factors in a particular field it has been seen is possible to achieve. What is needed is a very high degree of mental discipline and unwavering adherence, to meet the specified parameters. Increasing sophistication in science and technology leads to more and more closely structured organizations. These provide very little independence of action or thought to an individual. There is a known and inherent element in a human being which causes resistance to severe structuring. This is a pointer towards incidents/accidents where lack of flying discipline has been the primary cause. In advanced countries, technology evolution has been gradual accompanied by steadily improving results. Individuals in such a society could systematically acclimatize to the stress placed on them in the form of higher demands for mental discipline and the need to curb the general tendency or urge for independent action. In the developing countries such as ours, instead of gradual, technical evolution, there has been a technology explosion in the last two or three decades. We have been suddenly exposed to a high level of technology. The society or the individual had insufficient time to adjust to the constant, demands of high technology. This is a basic cause of aberrations in flying discipline. It has been much in evidence in the early 50's in the USAF despite about 300 years of backup of the growing technology. Later during the 60's by the royal Air Force. This is an important element, which must find a place in our training. It is incorrect to state that the pilot could throw around the aircraft around like a tempest or hurricane. He cannot afford to indulge this practice in even a very small way with high performance aircraft like Jaguar, Mirage 2000 and the F16 class. This aspect requires to be emphasized adequately in pilot training from the beginning. It is a continuous process and only though rigorous training can we achieve the very high degree of psycho motor skills and developmental discipline required to create a safe flying environment.

3.2.2 Effects of Sudden Stress

The Pilot is subjected to sudden stress when he needs to cope with the following situations:

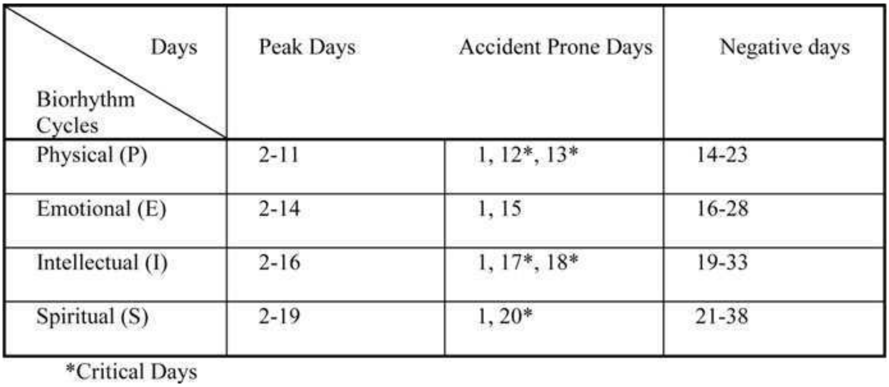
(A) Technical Defects in Aircraft: A pilot is subjected to sudden stress when some aircraft systems fail in the air. For example, failure of one engine of a twin engine aircraft or the failure of the under carriage system when it fails to come down according to need.

(B) Cognitive Factor: Cognitive factor arises during a conflict situation; it occurs when each pilot tries to establish identity and needs recognition. Cognitive conflict is found predominantly in the military aviation by the Air Force.

4 Behavioral Analysis and the Impact of Cycles of Biorhythm on Human Performance: Bio Analysis

The combined individual performance under a combination of the influences of each of four cycles consists of the combination of positive, negative and zero cycles for Physical, Sensitivity, Intellectual and Spiritual Cycles, resulting in behavior influence possibilities shown below. For days other than those critical days or for about 80% of the time the positive and negative positions of the cycles result in a change of the behavioral pattern as shown below:

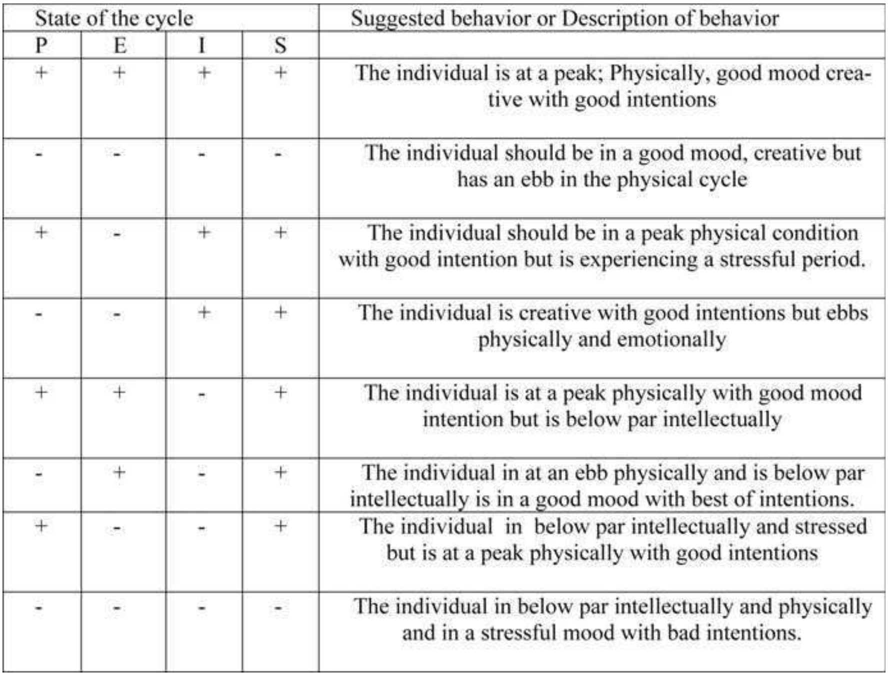
Table 1. Accident Prone Days



Days Biorhythm Cycles	Peak Days	Accident Prone Days	Negative days
Physical (P)	2-11	1, 12*, 13*	14-23
Emotional (E)	2-14	1, 15	16-28
Intellectual (I)	2-16	1, 17*, 18*	19-33
Spiritual (S)	2-19	1, 20*	21-38

*Critical Days

Table 2. Combination of Biorhythm behavior



State of the cycle				Suggested behavior or Description of behavior
P	E	I	S	
+	+	+	+	The individual is at a peak; Physically, good mood creative with good intentions
-	-	-	-	The individual should be in a good mood, creative but has an ebb in the physical cycle
+	-	+	+	The individual should be in a peak physical condition with good intention but is experiencing a stressful period.
-	-	+	+	The individual is creative with good intentions but ebbs physically and emotionally
+	+	-	+	The individual is at a peak physically with good mood intention but is below par intellectually
-	+	-	+	The individual in at an ebb physically and is below par intellectually is in a good mood with best of intentions.
+	-	-	+	The individual in below par intellectually and stressed but is at a peak physically with good intentions
-	-	-	-	The individual in below par intellectually and physically and in a stressful mood with bad intentions.

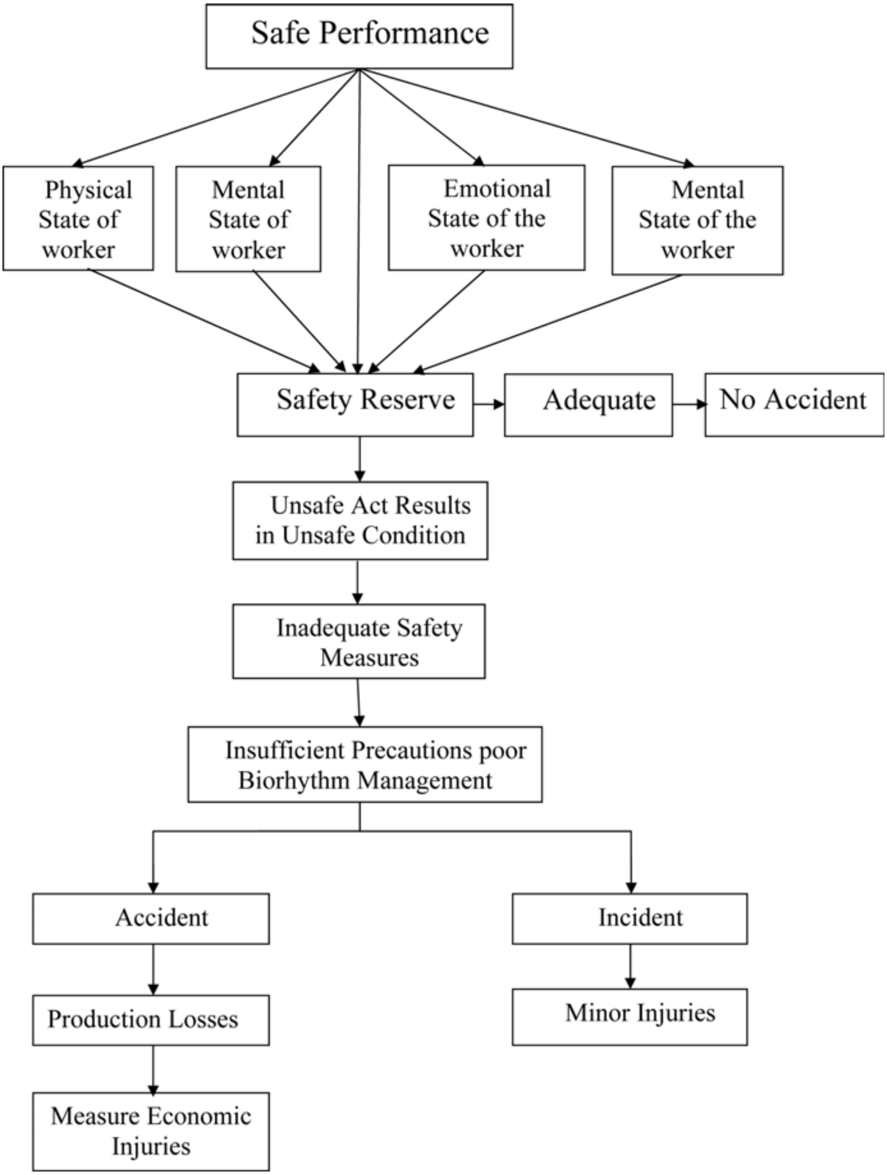


Fig. 2. Anatomy of Accident Flow Chart

4.1 Behavior Based Biorhythm Management

A good manager is a successful planner who strives to achieve numerous objectives by overcoming work losses or errors. It means that the managers will have managed to execute their plans to achieve their goals despite unplanned events occurring. If these unplanned events are on the increase it becomes necessary to reduce their effect by skillful management or by skillful use of the biorhythm. The essence of biorhythm management is to establish a set of procedures by which the biorhythmic accident prone period of each individual worker is considered and compensated for on an individual basis. In this way the likelihood of worker performance accidents may be minimized.

4.2 Causes of an Accidents and Safety Measures

An accident is an unplanned, uncontrolled, and undesirable event. It may be a sudden mishap which interrupts an activity or a function. A narrow miss which avoids an accident or a mishap is much the same. We know that the results of the accident frequently cause a delay in production, damage to material, and reduction of quality and possibly the loss of life.

Contributing Factors: These are frequently an accumulation of unsafe acts and unsafe conditions which provide circumstance predisposed to cause an accident. The Supervisor plays an important role in preventing accidents. Effective Safety Management necessitates eliminating and controlling these contributing factors.

Safety Responsibilities and Measures

- Recognize unsafe acts
- Training and monitoring staff
- Recognizing poor performance
- Enforcing Safety rules
- Planning for Safety as an essential part of the job
- Correcting and identifying hazards which are due to improper or inadequate maintenance
- Providing and enforcing clear Safety instructions
- Insuring Safety devices are adequate and maintained.
- Proper Overall Safety management

The worker is the second major contributing factor. The mental, physical, emotional and spiritual condition of the worker results in the following:

Mental Conditions:

- Possible lack of Safety awareness
- An improper attitude
- Inattention to the job at hand

- Slow mental reaction
- Poor management
- Substance abuse

Physical Conditions:

- Extreme fatigue
- Lack of coordination
- Physically unqualified for the job
- Loss of sleep
- Substance Abuse

Emotional Conditions:

- Lack of emotional stability
- Nervousness
- Temperament
- Loss of family members
- Under Heavy Stress environment
- Home life
- Lack of sleep

Immediate causes:

- Unsafe acts
- Unsafe condition
- Failure to use protective equipments or tools
- Hazardous movement
- Inadequate illumination
- Poor ventilation
- Bad house keeping
- Ineffective or Inadequate Safety devices

Safety measures:

- Develop Safety awareness and Safety consciousness
- Adequate training
- Proper Safety environment
- Use of Biorhythm model for “off days”
- Precautions against Accident Prone days
- Improving worker morale
- Improving attitude
- Improving working environment
- Managing stress
- Avoid Fatigue
- Avoidance of bad habits
- Provision of tangible rewards for accident free periods

Identifying the primary factors which influence safety is a complex task and no single factor can be an adequate predictor of a worker's or supervisor's performance or susceptibility towards making and reducing errors. Biorhythms provide a potential management tool to inform and aid better decision making by a consideration of the periods when a worker may be most susceptible to errors.

5 Analysis of the Causes of Human Error Factor and the Development of a Model for Predicting Human Error

Error may be considered in its most basic form as an intended action which is not correctly executed. At a simple but often opaque level errors are actions which result from various unconscious or conscious mistakes. At the most transparent level an error may be considered as any act which results in an accident or incident that may involve personal injury, death and damage. Studies show that human error rates doing simple repetitive tasks can normally be expected to occur about once in 100 events. It has also been demonstrated that under certain circumstances human reliability can be improved by several orders of magnitude. An error rate of 1 in 1000 is regarded as good. The British Civil Aviation Authority for instance requires that automatic landing equipment must not suffer a catastrophic failure more than once in 10 million landings. Human errors vary widely depending on the task and many other factors such as Physiological factors.

> ***Fatigue:*** Fatigue can be defined as The Level of Reduced Performance in which there is no certainty that a person can react in an emergency even when the need is obvious. Long working hours and inadequate rest periods are potential threats as they result in inadequate responses.
>
> ***Sleeplessness:*** It adversely affects and lowers the performance due to a lack of focus. It is difficult to function at an adequate attention level.
>
> ***Lack of motivation:*** It has been noticed that in many accidents that a well trained and highly capable pilot or technician has made an error. In many examples it is most likely due to a low motivation level. Selection, training and periodic checking can help ensure the ability to perform a given task for a given period of time. The continuing performance of a person showing the same dedication is largely governed by his motivational level and not by his ability.

> ***Stress:*** Extended working hours usually cause a person stress. All personnel are trained and conditioned to cope with this situation. However, if the stress levels increase drastically due to totally unrelated factors such as, death of a spouse, administrative problems, financial problems, the likelihood of a person making an error increases dramatically. Misunderstanding and lack of information is part of work culture. It can be reduced or contained by the use of concerted efforts.

5.1 Man Machine Interaction When Considering Flying Mistakes

A Communications study by Michaela A. and Dorheim Moffect shows that in 107 reports involving competing tasks, the biggest cause was distraction, 68 of 107 incidents were due to communications problems. Crew members who communicate well tend to perform better but conversation demands attention. For example it is necessary to think of responses and retain them in memory until it is your turn to speak. Flying crew may become preoccupied with the conversation and not notice other demands on their attention. Head down activity (22 out of 107) such as the performing the flight management system or re viewing approach charts, was a factor in 22 incidents. When engaged in head down activities the eyes are diverted and can require a high degree of concentration. When the head is down the pilot cannot reliably monitor the flying during longer tasks. When the head is down it is necessary to suspend programming in mid stream and 13 out of 22 such cases caused a failure by the pilot to monitor flying. Response to an abnormal situation was involved in 19 out of 107 incidents. In one example, the crew had to deviate around thunder storms and then had to descend and the cabin pressurization slowly failed. They had forgotten to reset their altimeters when descending to the lower air space and were set at 300 feet low at an attitude of 13000 feet. Treating an urgent situation narrows the focus of attention. Visually searching for other traffic was a factor in 11 incidents out of 107. One air crew was reported as receiving an alert and then missing the turn at an intersection because they were looking for other traffic. Head down work when searching requires attention and it takes the eyes from where they should be. In all of the above cases the main routine task was neglected. The neglected duties were monitoring the status of the aircraft and the pilots flying. Failure to correctly monitor occurred in 69% of cases. Altitude Deviation was the neglected task in 31 of the 107 cases. Sixteen of these cases involved not making a required turn and 13 of the cases involved a failure to reset the altimeter. There were 107 reports. The percentage failure probability of making flying mistakes considering various cases has been plotted on the Y-Axis of figure: 3.

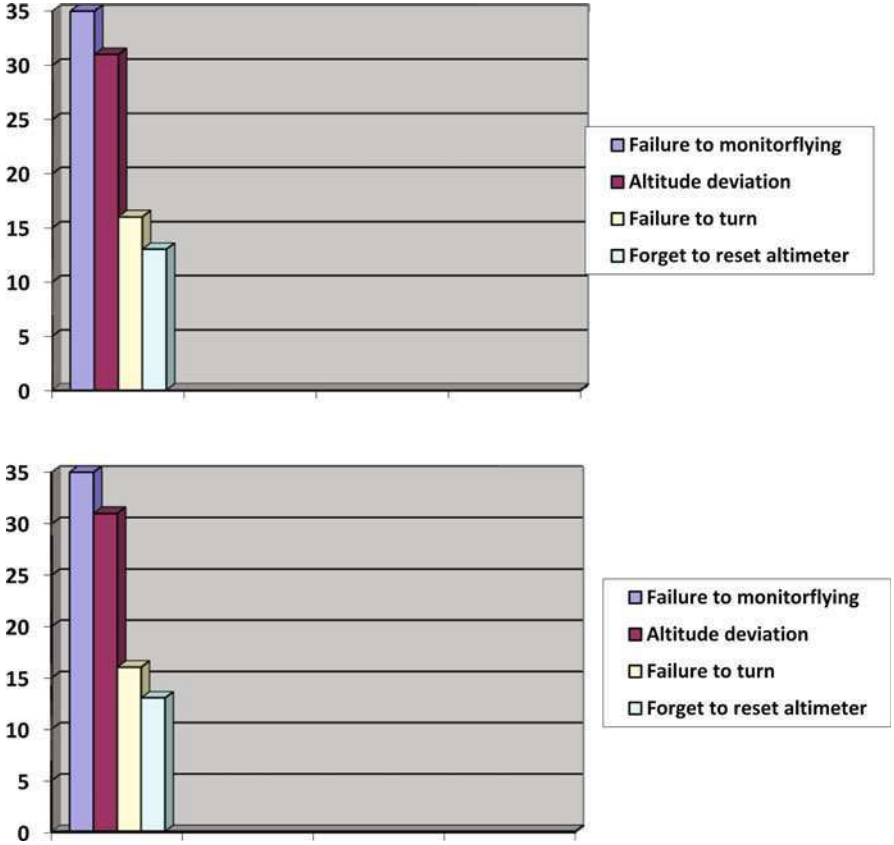


Fig. 3. Man Machine Interaction and Flying Mistakes

5.2 An Error Taxonomic Approach

An analysis of 93 major accidents over a period of 24 years from 1959 to 1983 has been analyzed. Sears 1986 found that 12% which is 11 accidents were caused by maintenance and inspection deficiencies. Similarly Nagel 1988 reports that four out of every hundred accidents that occurred from 1977 to 1988 were the result of maintenance error. During an aircraft inspection certain defects were found in an aircraft which was ready to fly. This was clearly unacceptable. It is pertinent to make this requirement explicit by providing information about the requirements. These requirements are based on human error avoidance.

5.3 Factors Which Affect Decision Making

Decision making is the task during which any potential defect is located by searching and evaluated in order to determine whether it should be reported. In this task both Type I errors (False Alarms) and Type 2 errors failure to FIN can occur. These errors have their own “Tradeoff Relationship”. Some combined accuracy measure must be derived before any tradeoff between the search Speed and the Accuracy can be determined.

A particular model of the human being having the ability to provide a rational economic maximization which has received widespread support for inspection is Signal Detection Theory (SDT), it was originally proposed by Swets and co-workers (e.g. Swets, 1967). This was to serve as a model of how humans are able to detect Signals in the presence of noise. It was subsequently applied successfully to the inspection problem by Wallack and Adams, 1969, 1970; Sheehan and Drury, 1971; Drury and Addison. 1973.

In the SDT the inspector is assumed to be making a choice for each item inspected as to whether the item contains a defect (“signal”) or does not (“noise”) As the evidence for signal or noise is somewhat equivocal, there is assumed to be an “Evidence Variable”. This increases when a signal is present and equivocal. There is also assumed to be an “Evidence Variable” which increases when a signal is present and decreases when only noise is present. An example of its use would be the judgment whether a dent in a stabilizer leading edge should be reported. Dents can range from almost imperceptible to those which must be reported. The Evidence Variable (dent visual severity) must be judged against both Written Size Standards and the likely effects of the dent or the Flight Characteristics.

5.4 Factors Affecting Sensitivity

Most factors affecting discriminability or sensitivity are physical. They can be characterized as the perceived difference between the observed indication and the relevant standard. Thus, indications obviously well above or below the standard will have high discriminability (d) values. Examples would be the existence of large areas of corrosion, cracks which are noticeably larger than those allowed or completely missing rivets. None would require difficult or Error Prone decisions. The expression “perceived difference” implies both High Signal and Low Noise in SDT terminology. Low noise means low levels of visual distraction. That is competent cleaning, Low levels of fatigue that is frequent task breaks; Very Clear Standards well defined and well presented job aids. All of these can be improved for the aircraft inspection task. Comparison standards at the work place have been shown to be effective in improving discriminability. It should be possible for the inspector to make a direct side-by-side comparison of fault indication by using a standard. For example, the critical amount of corrosion beyond which a report must be made should be indicated by a life-sized diagram on the work card. If different corrosion types are present, life-sized photographs help in reaching a positive identification.

5.5 Factors Affecting Criterion

From SDT, the factors affecting the choice of criterion are the relative costs of making errors such as Misses and False Alarms and the True Rate of Defects (p). Using these factors, the optimum criterion can be calculated. This is rarely the exact criterion which is used by the inspector. In laboratory tasks and in non-aviation inspection tasks the inspectors will choose a criterion in a conservative manner. For examples, if the criterion should be low they should be very willing to report indications as defects. Inspectors will choose a criterion which is not low enough. Similarly, they will choose a criterion which will not react quickly enough in changing their criterion as the costs and probabilities will change. It is important to provide accurate and up-to-date feed forward information on the probabilities of defects in different areas in order to allow the inspector to make rapid criterion changes.

There are also known criterion shifts that consider both the changing defect rate and the time spent on the task. There is little to be done when increasing the defect rate. It is fixed by the state of the aircraft. The reduction in the hit rate at very low defect rates may set a limit to the use of humans as detectors of rare events. Paradoxically, as the maintenance improves and gives fewer defects, the ability of the inspector to detect the few remaining defects worsens. There is a need for more research into the human or machine function allocation to alleviate the low defect rate problem. The time on the task, the vigilance phenomenon, cause a reduced detection rate due to the Criterion Shift under special circumstances. That is uninterrupted performance. This may not be a problem of aircraft inspection, although heavy use of night shift inspection where interruptions are less frequent and the human less vigilant, requires further study.

5.6 Rationale for Research on Visual Inspection Training

From the above discussion, training for visual search would be expected to result in reduced search errors. That is type 2 errors and a reduced search time. Similarly, training for decision making and perception can be expected to result in reduced type 1 and type 2 errors. Although training can be used to improve visual inspection performance, specific training schemes are not associated with factors that determine improvement in visual inspection performance. Hence, training schemes are developed that guarantee improvements for a particular task without consideration as to whether such a training scheme could be extended to a similar task, a different task, or whether the training is optimizing the use of instructor and the trainee time. The first step in the development of a rational training scheme is to identify the factors that affect visual inspection performance. The next step is to determine which of the functions of the inspection task can be developed by training. This in turn will serve to establish the sensitivity of the inspection parameters for training.

For any training scheme to be effective it should minimize both Search Errors and Decision Errors. Referring to the earlier proposed model of visual inspection, it is observed that intervention strategies could be developed at various stages of the inspection process. This could be hypothesized to change the inspection parameters to achieve an improved performance.

The following factors are critical in the search process:

(a) The ability to identify salient features which can be associated with a particular defect. This is so that features may be searched for in parallel instead of requiring separate attention.

(b) Visual search
(c) Eye movement scanning strategy.

In order to improve the visual inspection performance it is necessary to develop training schemes which provide improvements in the above factors. The following section briefly describes various training schemes.

5.6.1 Visual Lobe Training

The visual lobe is a very important determinant on search performance. Those observers with a larger visual lobe require fewer fixations than observers with a smaller visual lobe. A large visual lobe or peripheral acuity may account for superior search performance. We still need to know how a large visual lobe can affect search performance and how people can be trained so as to increase the size of the visual lobe. If the above questions could be answered, this could result in a strategy to improve the visual lobe. More general questions then arise. How does lobe size training generalize across tasks such as targets and back grounds. We wish to understand whether the visual lobe training on a given target; would result in an improved search performance for a different target type and the sensitivity of the appropriate search parameter for this type of training. It is essential to identify whether a cross-over effect exists. If it does, then it is sufficient to train only on one target type. If not, then it is necessary to identity various target subsets, for examples T_1, T_2, where the cross-over occurs. Trainees could then be provided visual lobe training on a single target of each target subset.

5.6.2 Feedback Training

A person needs rapid accurate feedback to correctly classify a defeat, or the effectiveness of a search strategy. Every training program should begin with frequent feedback and to gradually reduce this until the level of proficiency required is reached. Additional feedback beyond the training program will help to keep the inspector “calibrated”.

The following feedback could be provided:

(i) Feedback showing accuracy of classifying defective items into the correct categories
(ii) Feedback of the search strategy derived from monitoring eye movements
(iii) Feedback of fixation times from the subjects eye movement search.

The First item is known to be essential in learning perceptual tasks (Annett, 1966). It provides the novice with information regarding a critical difference between a defective item and the satisfactory item. This helps to develop a mental attitude which contains the internal characteristics of a defective item. We are, however, still unsure as to what has been improved. For example, learning resulted in producing a new internal conceptual model of the task or is the inspector using only certain dimensions of the fault to classify it.

It has been shown that an important difference between the best and the poorest search performance is the length of the sweeps between eye fixations during the search task. Does there exist a difference between how a novice and an expert move their eyes across the visual field. Gould (1973) during a visual inspection study of circuit chips found that most of eye fixations occur within a definite boundary, which is the area most likely to contain the targets. It has been demonstrated that eye movements during a visual search occur based on knowledge of the location of faults and on the probability of them occurring. The question that needs to be answered is: Does feedback information regarding the eye movements help to improve the scanning strategy? We hypothesize that providing such feedback information would aid the inspectors by allowing them to identify areas which have been not covered or areas where excessive time is spent. By helping the inspector to develop such a strategy it will become possible to cover the entire area more effectively.

5.6.3 Feed Forward Training

When a novice inspector with no knowledge of the type of faults, the probability of faults, and the occurrence of faults, conducts a visual search, it would be expected to be inefficient. Providing feed forward information should result in an improved search strategy. This is because the uncertainty is reduced as the inspector knows both where to look and what to look for. Perhaps the inspector may use the information to achieve a more systematic search strategy, guided by knowledge of the fault characteristics. The inspector could use feed forward information in the following ways:

(1) To ignore the information completely
(2) To selectively incorporate some of the information
(3) To incorporate this information only at later stages of inspection,

That is only after gaining some verification. Using this suggests that experienced inspector's make use of feed forward information that complements their existing sensitivity to the fault. If the fault is one that is not easily detected, then the inspector relies heavily on any information provided. Inspection tasks that will most likely benefit from this addition to prior information include those where the value of inspection time is great. Those faults have been looked easily in which the fault is particularly difficult to detect and those in which the product may contain rare detrimental.

5.6.4 Attribute Training

Consider an item A. Let the item be faulty on attributes A1, A2, A3 and A4. The inspector could be trained on each of the above attributes such training would allow the inspector to set a response criterion for each attribute. The training should be generalized in the sense that the inspector should be able to classify the items as defective if the items are faulty on one or more of the attributes. Main attributes of the systems reliability are probability of the system, performance of functions adequately, operating conditions and time periods (Saket, R.K.; Wg. Cdr. Kaushik, S.P. and Col. Singh, Gurmit, 2008). The inspector could be trained on which of the attributes match. Firstly based on the probability of the item being faulty on these attributes and the ease with which the matching occurs. Experience and training of the inspectors determine how defective attributes are arranged (Goldberg and Gibson, 1986).

5.6.5 Schema Training

It is essential that the subject develops an appropriate valid mental template or internal representation schema of the fault. The key to the development of such a schema is that it should provide for successful extrapolation for use in novel situations which are still recognizable instances of the schema. We need to know how schemas are developed. Whether inspectors can be trained to develop schemas, and what sort of training being either Rule based or Knowledge based must be provided to the inspectors for effective development of such schemas. The effects of the two methods of training need to be evaluated during schema development: “active training” and “passive training”. In active training, the inspector is presented with various instances of fault and no-fault. He has to classify them as defective or non-defective. Feedback is provided regarding the correctness of this classification. In contrast, passive training is where the inspector is merely presented with various instances of the faults without being required to provide an active response.

5.6.6 Interaction of SRK Behavior

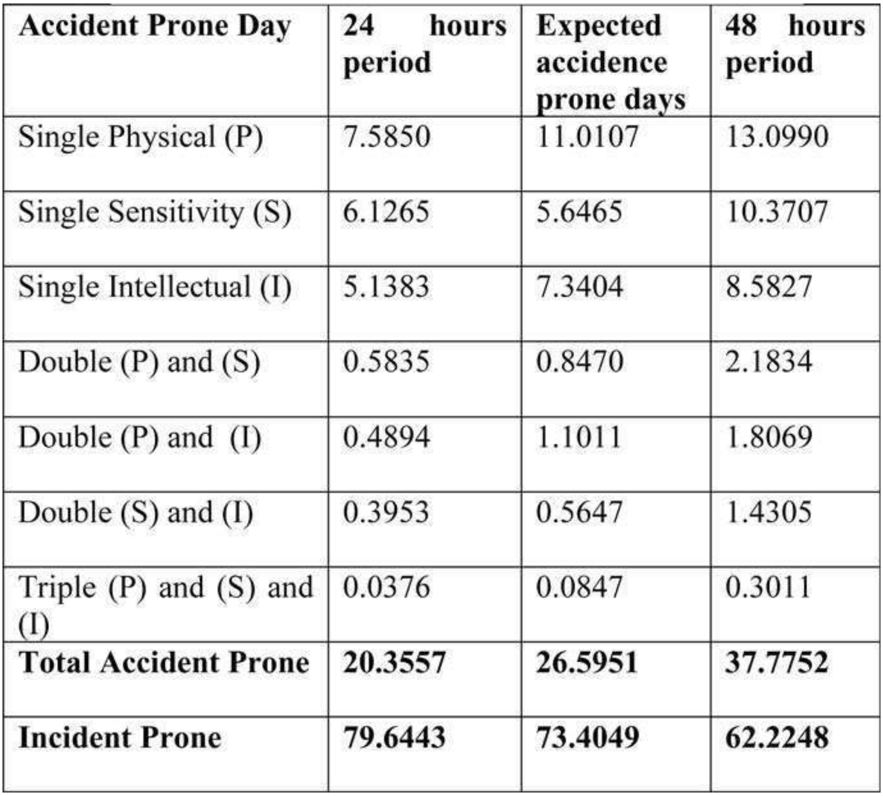
It must be explained at this point that in aircraft inspection Skill Based (S), Rule Based (R) and Knowledge Based (K) behaviors are rarely stand alone or distance behaviors modes, indeed they overlap on some occasions and support each other on others. For example the skill based behavior of probe movement is supported by either knowledge based for line choice or rule based for expert behavior that exists on the boundaries of the movement. The probe should not cut the overhead line and a movement too close to an edge should be avoided. Similarly, rule based behavior and clarification in visual inspection is sometimes superseded by knowledge based behavior that is based on active reasoning on a deeper level and a functional understanding of the aircraft. During Virtual Inspection of line wing leading edge, the inspector who is looking for dents may realize that a dent forward of another dent may be more important because it could cause problems in fight control. This and the preceding example highlight control and the often symbiotic relationship of default behavior.

6 Statistical Technique for Biorhythmic Analysis

In any Biorhythm research analysis it is necessary to scientifically demonstrate whether or not a relationship exists between the biorhythm and the human performance (Douglas, Neil and Francis, L Sink (1976). Hence fundamental concepts in probability and statistical analysis have been applied to use the Gaussian distribution approach for biorhythmic analysis used to prevent aviation accidents. Biorhythm is reviewed in this paper. The probability of occurrence of accident-prone days is 21.9%. This value is reduced to 20.4% by excluding multiple critical days.

The probability of occurrence of single, double and triple critical days has been shown in following table 1. It is to be noted that the longer the cycle period the higher the probability of a zero crossing. Since the accident-prone periods for the physical and intellectual cycles will alternately centre on midnight and noon for the adjacent cysts, it is often difficult to accurately assign an accident to this Biorhythm period. It is therefore convenient to assess a 48 hrs period for alternate periods of the physical and intellectual cycles. The expected percentages of occurrence of accident-prone days are presented in column 3 of the table. It is found that in this case the probability of occurrence of accident-prone day is 26.6% each biorhythm accident-prone period may also be analyzed as a 48 hrs period. The 4th column of table shows the expected percentage of occurrences of accident-prone periods for a 48 hrs period. The expected occurrence of these accidents prone periods is 37.8%.

Table 3. Accident Prone Days (percentage)



Accident Prone Day	**24 hours period**	**Expected accidence prone days**	**48 hours period**
Single Physical (P)	7.5850	11.0107	13.0990
Single Sensitivity (S)	6.1265	5.6465	10.3707
Single Intellectual (I)	5.1383	7.3404	8.5827
Double (P) and (S)	0.5835	0.8470	2.1834
Double (P) and (I)	0.4894	1.1011	1.8069
Double (S) and (I)	0.3953	0.5647	1.4305
Triple (P) and (S) and (I)	0.0376	0.0847	0.3011
Total Accident Prone	**20.3557**	**26.5951**	**37.7752**
Incident Prone	**79.6443**	**73.4049**	**62.2248**

Carlos M. Perez (2004) found that four principles are essential for successful hostage rescue mission: surprise, intelligence, operation's skill, and deception. These principles are derived from planning models used by special operations, personal experience, and an analysis of six historical case studies. The normal distribution curve based on data available in air force has been used; the area under normal distribution curve represents the accident zone. In this example the accident zone is that area of operation during which the performance demand exceeds the performance ability of the aircraft pilot. The Incident-duration curve of aircraft accidents has been designed by Saket, R.K. and Kaushik, (Wg. Cdr.) S.P. (2005). The incident-duration curve has been assumed to be a straight line and increases with the duration for the reliability evaluation of the aircraft system.

7 Reliability Evaluation of Biorhythmic Aviation Accidents

Reliability of the aviation system is defined as the overall ability of the aircraft system to perform its function. Reliability theory as an extension of probability theory was first applied in electronics, nuclear and space industries after World War-II, where high reliability was a requirement for these increasingly complex systems. Nowadays reliability studies are performed in almost all engineering branches. This chapter presents a new methodology for the evaluation of the probability of an accident (P_{ac}) based on Biorhythmic approach. The methodology has been developed in this paper using normal distribution curves of the accidents. Biorhythmic accidents have a continuous distribution function and have a Gaussian distribution for a specified time interval.

7.1 Gaussian Distribution Approach

The performance demand model for pilots flying has been a Gaussian distribution for a specified time interval according to the incident – duration curve [Arya L. D., Chaube, S. C. and Saket, R .K. (2001)].

$$f\left(P_d\right) = \frac{1}{\sigma_d\sqrt{2\pi}} e^{-0.5\left(\frac{P_d-\overline{P_d}}{\sigma_d}\right)^2} \tag{1}$$

The aggregated performance capacity model of the pilot and the air craft system has been approximated as Gaussian.

$$f(C) = \frac{1}{\sigma_c\sqrt{2\pi}} e^{-0.5\left[\frac{C-\overline{C}}{\sigma_c}\right]^2} \tag{2}$$

The failure probability (P_F) of the above performance demand and capacity models can be written as follows:

$$P_F = (1-P_s) \tag{3}$$

The success probability of the model (P_s) can express as follows:

$$P_s = \int_{-\infty}^{\infty} \int_{-\infty}^{C} \frac{1}{2\pi\pi_c\sigma_d} e^{-0.5\left[\frac{C-\bar{C}}{\sigma_c}\right]^2} e^{-0.5\left(\frac{P_d-\bar{P}_d}{\sigma_d}\right)^2} dcdP_d \tag{4}$$

After substitutions, equation (4) can be written as follows.

$$P_s = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z} e^{-0.5\left(x^2+y^2\right)} dy \right] dx \tag{5}$$

In the view of the above substitution (5) the success probability of the system can be written as follows using Gaussian distribution approach [Saket, R.K.; Wg. Cdr. Kaushik, S.P. and Col. Singh, Gurmit (2008)].

$$P_s = \int_{-\infty}^{\infty} \int_{-\infty}^{\beta} \frac{1}{2\pi} e^{-0.5\left(x'^2+y'^2\right)} dy'dx' \tag{6}$$

$$\text{Where, } \beta = \frac{\bar{C} - \bar{P}_d}{\sqrt{\sigma_d^2 + \sigma_C^2}} \tag{7}$$

The limit β comes out to be independent of x′. Further equation (6) can simplified as follows:

$$P_s = \int_{-\infty}^{\beta} \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-0.5\left(x'^2\right)} dx' \right] e^{-0.5\left(y'^2\right)} dy'$$

$$= \int_{-\infty}^{+\beta} \frac{1}{\sqrt{2\pi}} e^{-0.5\left(y'^2\right)} dy' = \varphi(\beta) \tag{8}$$

In fact φ (β) is the success probability of the air craft system and represents the area under the normal distribution curve having mean zero and standard deviation $\int [N(0,1)] = 1$. Various curves based on equation (8) have been plotted using MATLAB simulation. This expression satisfies the Gaussian distribution approach to the pilot's reliability evaluation. φ (β) is the area under the normal distribution curve having zero mean and standard deviation [N (0, 1)] is one from - ∞ to β. This value can be conveniently obtained from standard tabulated data. Daily variations in the mental load or performance demand on the pilot can be accounted for by predicting the various demand levels P_{di} and the relative frequency of accident occurrence of these levels are assumed to be Lo, L1, L2 ………, Li. The frequency of occurrence is $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_i$. For each demand level probability of failure can be calculated and overall probability of failure is given as

$$P_f = \sum_i \alpha_i P_{fi} \tag{9}$$

Various plots of P_f v/s $\frac{\overline{C}}{P_d}$ have been plotted in this paper. The curves shown here can be used as a standard curve for evaluating the pilot's capacity.

7.2 Safety Factor Concept and Peak Demand Considerations

The probability distribution function of the pilot capacity has been earlier obtained as a Gaussian. Further, it has been determined that the peak performance demand dominates over low-level loading, whereas, the probability of failure under low load level condition is negligible. P_{dmax} is the peak performance loading / demand on the pilot. The safety factor 'S' is defined as:

$$S = C/P_{dmax} \tag{10}$$

It is obvious that the pilot's ability to fly / flying capacity 'C' has a normal distribution and S is a random variable. Since P_{dmax} has been considered to be constant, the safety distribution function 'S' will also be a normal and is given as follows [Arya L. D., Chaube, S. C. and Saket, R .K. (2001)].

$$f_s = \frac{P_{dmax}}{\sqrt{2\pi}\,\sigma_c} e^{-0.5\left(\frac{P_{dmax}\ S - \overline{C}}{\sigma_c}\right)} \tag{11}$$

The mean safety factor and standard deviation of the safety factor is given as:

$$\overline{S} = \frac{C}{P_{dmax}} \quad \text{and} \quad \sigma_s = \frac{\sigma_c}{P_{dmax}} \tag{12}$$

The probability of the performance failure is given as under:

$$P_F = \int_{-\infty}^{1} \frac{P_{d\max}}{\sqrt{2\pi}\,\sigma_s} e^{-0.5\left(\frac{s - \bar{c}/P d\max}{\sigma_s}\right)^2} ds = \phi \left[\frac{1 - \dfrac{\overline{C}}{P_{d\,\max}}}{\dfrac{\sigma_c}{P_{d\,\max}}} \right] \tag{13}$$

The Failure Probability P_F v/s $\dfrac{\overline{C}}{P_{d\max}}$ curves have been plotted using MATLAB simulation. There are the standard curves available for evaluating the performance capacity of the pilot based on the safety factor concept and peak load considerations [Saket, R.K.; Wg. Cdr. Kaushik, S.P. and Col. Singh, Gurmit (2008)] .

7.3 Performance Evaluation Based on Peak Demand Using Simpson 1/3rd Rule

The LOLP is one of the most commonly used indexes for planning the performance capacity of the pilot. This index is generally obtained by convolving the performance model with a demand model. All types of composite reliability indices such as the loss of performance demand probability, and accident frequency have been assessed not only for the overall aircraft system but also for single components and aircraft pilots. Failure probability has been evaluated using a more realistic model as Incident - Duration Curve. A Stepped Incident-Duration Curve has been considered for aircraft system reliability evaluation using Simpson's 1/3rd rule. In the following expression, 100 small steps have been considered in Daily Incidence - Duration Curve. The performance model adopted is a Normal Distribution Function and the evaluation is based on the Maximum Average Performance Capacity of pilot available. The probability of the performance demand exceeding the performance capacity (LOLP) of the pilot using a Stepped Incidence - Duration Curve can be written as follows:

$$\text{LOLP} = \int_{0}^{100} \frac{t}{100} \left[\int_{-\infty}^{P_d(t)} \frac{1}{\sqrt{2\pi}\sigma_c} e^{-0.5\left(\frac{C-\overline{C}}{\sigma_c}\right)^2} dc \right] dt \tag{14}$$

Putting, $\dfrac{C - \overline{C}}{\sigma_C} = Z$

The LOLP expression of (14) can be expressed as:

$$\mathrm{LOLP} = \int_{0}^{100} \frac{t}{100}\left[\int_{-\infty}^{P_d(t)} \frac{1}{\sqrt{2\pi}\sigma_c} e^{-0.5(Z)^2} dz\right] dt$$

$$= \int_{0}^{100} \frac{t}{100}\, \varphi\left(\frac{P_d(t) - \overline{C}}{\sigma_c}\right) dt \qquad (15)$$

The LOLP of the composite aircraft system can be evaluated by using the above methods of area evaluation for any step of the Incident - Duration Curve [Arya L. D., Chaube, S. C. and Saket, R .K. (2001)]. The Simpson's 1/3rd rule has been used to evaluate the LOLP of the pilot of the aircraft system. Reliability of the air craft operation or the success or failure probability of the aircraft system based on biorhythm theory has been evaluated using Simpson's 1/3rd rule considering small time steps of durations and various operation periods of aircraft systems.

8 Results and Discussion

The failure probability due to the biorhythmic effect on the pilot of the aircraft system which has been explained earlier was evaluated by assuming the following available data. The mean performance ability of pilot to fly (P_a) in percentage and standard deviation data are give in following table (4).

Table 4. Mean performance ability of pilot (P_a) and standard deviation data

(Pa)	90	89	88	87	86	85	84	83	82	81	81.5	80.1
σa= 5%	4.50	4.45	4.40	4.35	4.30	4.25	4.20	4.15	4.1	4.05	4.02	4.005
σa=10 %	9.0	8.9	8.8	8.7	8.6	8.5	8.4	8.3	8.2	8.1	8.05	8.01
σa=15 %	13.5	13.35	13.2	13.05	12.9	12.75	12.6	12.45	12.30	12.15	12.07	12.02

The probability of the accident or failure probability of the Air Craft Pilot's ability to fly due to biorhythmic effects at mean performance demand (P_d) 80% has been evaluated using Equation (9). The biorhythmic accident probability (P_{ac}) at different conditions taking account of Pilots ability to fly is given in table (5).

The graphs of the probability of accident by aircraft pilot due to biorhythm and mean performance ability are shown in fig. (6) and (7). The Distribution Function for both the performance ability to fly and the performance demand on the aircraft pilot has been adjusted to be normal. From fig (6), the performance demand of the aviation system is constant at 80% and performance capability of pilot increases gradually. At different performance abilities of the pilot after considering biorhythmic effects the failure probability of system has been evaluated. The failure probability of the aircraft due to biorhythmic effect of the pilot decreases with the increase in the air craft pilot's capability. When the performance ability and demand are equal the success or failure probability is equal to 50%. If performance demand exceeds the performance capability of the pilot due to biorhythm, there is 100% probability or certainty of an aviation accident.

Table 5. Pilots ability to fly and standard deviations

Pa %	Pac at σa= 10% σd= 5%	Pac at σa= 10% σd= 5%	Pac at σa= 10% σd= 5%
90.00	0.5485	0.6539	0.7637
89.00	0.5668	0.6788	0.7810
88.00	0.5901	0.7033	0.8015
87.00	0.6091	0.7327	0.8301
86.00	0.6539	0.7643	0.8483
85.00	0.6949	0.7981	0.8707
84.00	0.7451	0.8336	0.8897
83.00	0.8015	0.8707	0.9207
82.00	0.8632	0.9091	0.9443
81.00	0.9364	0.9483	0.9721
81.50	0.9641	0.9801	0.9880
80.10	0.9923	1.0	1.0000

The Failure Probability of the aircraft system due to biorhythmic effects has been evaluated using the safety factor concept and is shown in fig (6). At a constant performance demand of the aircraft system the safety factor increases with the performance ability of the pilot. This graph indicates that the Failure Probability decreases with an increasing value of the Performance Safety Factor. At constant Performance Ability and increasing Performance Demand, Failure Probability of the aviation system increases and the curves cut each other at 0.5. At this point, the success and failure probabilities become equal to 50%. If the demand increases with biorhythmic ability of the pilot, no one can prevent the aircraft system from an accident.

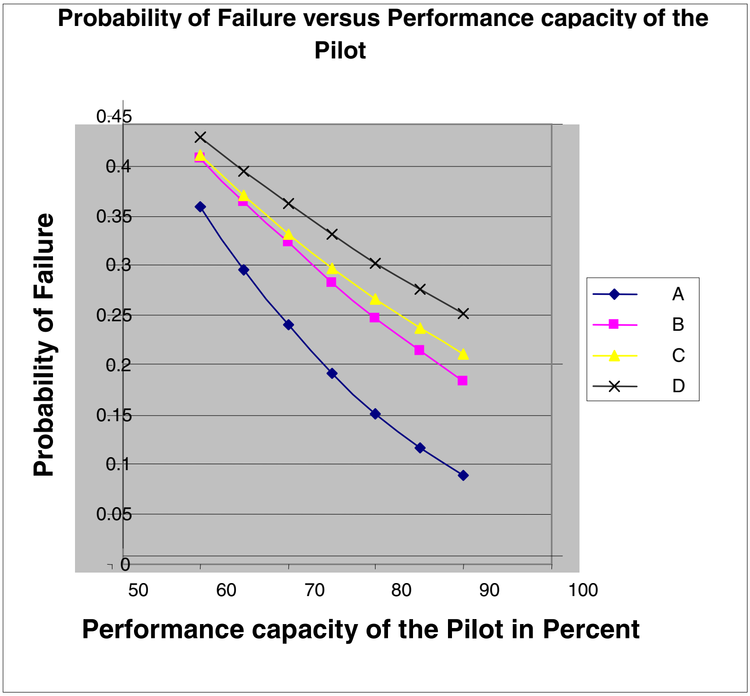


Fig. 6. Failure Probability (Aviation Accident Chances) versus Performance Capability of the Aircraft Pilot

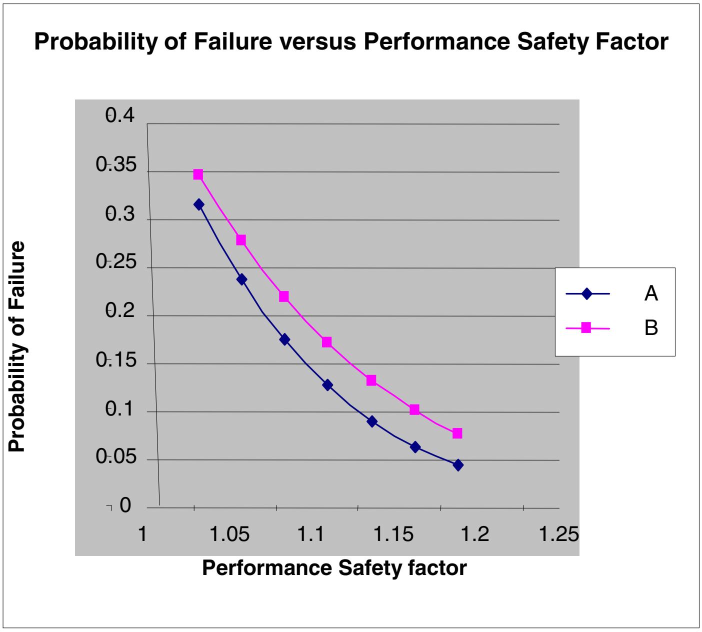


Fig. 7. Failure Probability (Aviation Accident Chances) versus Performance Safety Factor of the Pilot

9 Conclusion

It has been demonstrated that whenever pilot demand exceeds the Pilot Performance Ability, the probability of an air craft accident increases. Biorhythm plays a vital role in increasing the internal demands on the pilot and simultaneously diminishes the Pilots ability. The failure probability of the system due to pilot inability has been described by various plots as shown in the figures (6) and (7). The Failure Probability of the aircraft system decreases as the standard deviations of the performance ability of the pilot. Serious aircraft accidents at the rate of 70 to 80% are associated with the influence of human error. These are considered to be the most critical when the functional systems of aircraft fail and the pilot is exposed to adverse factors. Because a pilot's Psyche and physiological factors affect his performance in the air, errors are likely. Among the factors which lead to disturbances in pilot's psycho may be personal living experiences and events; various psychological factors, biorhythmic effects, reaction to emergency situations, effects of trainings, machine factors which are largely technical in nature and environmental conditions. It is necessary to study the physical cycle, the emotional cycle and the intellectual cycle to determine when a pilot is prone to an accident. Selection of pilot's who are not prone to accidents and who can impart training. It is necessary to determine whether or not the pilot can keep his cool during emergencies. Study of the behavioral aspects, the man-machine interface and reliability improvements will help to reduce the number of aviation accidents. Reliability analysis of the aviation system considering the biorhythmic effects according to typical biorhythm cycles at constant performance ability of the pilot. A variable biorhythmic demand has also been proposed for future research.

Acknowledgements

Authors are grateful to Aviation Research Centre and Maintenance Safety Inspection division of the Indian Air Force for providing the necessary facilities and with financial support to complete this chapter. The corresponding author dedicates this chapter to sweet memories of his great father Shi Ram Das Saket Vardhan who has passed away on August 17, 2008. Without his blessings, constant support and timely encouragement it would not have been possible to complete this work in its present shape. He would also like thank to Aanchal, Aakanksha, Siddharth and newly born son Siddhant Vardhan (May 15, 2009) for providing Biorhythmic Emotional and Intellectual supports during this first work for Springer-Verlag, Germany.

References

Arya, L.D., Chaube, S.C., Saket, R.K.: Generation System Adequacy Evaluation Using Probability Theory. Journal of The Institution of Engineers (India) 81, 170–174 (2001), `http://www.ieindia.org`

Perez, C.M.: Anatomy of a Rescue: What Makes Hostage Rescue Operations Successful? Master's Thesis, Naval Postgraduate School, Monterey, Canada, No: A396724 (September 2004)

Carvey, D.W., Nibler, R.G.: Biorhythmic Cycles and the Incidence Of Industrial Accidents. Personnel Psychology 30(3), 447–454 (2006), `http://www3.interscience.wiley.com/journal`

Sacher, D.: The Influence of Biorhythmic Criticality on Aircraft Mishaps. Naval Postgraduate School, Monterey Calif., AD0718387 (September 1974), `http://www.oai.dtic.mil/oai`

Douglas, N., Francis, L.S.: Laboratory Investigation and Biorhythm. Hypothesis Aviation, Space and Environment Medicine, 425–429 (1976)

England, C.E., Naitoh, P.: An Attempted Validation Study of the Birth Date-Based Biorhythm. Hypothesis Aviation, Space and Environmental Medicine, 583–590 (1980)

Meyer, G.E.: The Application of Probability Calculations For Bird-Aircraft Strike Analysis and Predictions Using Radar. Military Aircraft Operations, Air Force Weapons Lab, Kirtland, AFB NM (March 1974), `http://www.stormingmedia.us/10/1098/a109800.html`

Hines, T.M.: A Comprehensive Review of Biorhythm Theory. Reprinted From: Psychological Reports, Psychology Department; Pace University (August 1998)

Aldrin, J.C., Medina, E.A., Allwine, D.A., Ahmed, M.Q., Fisher, J., Knopp, J.S., Lindgren, E.A.: Probabilistic Risk Assessment: Impact Of Human Factors On Nondestructive Evaluation And Sensor Degradation On Structural Health Monitoring. Human Factors Engineering And Man Machine System, A876464, No: 3361503C5226 (September 2006), `http://www.stormingmedia.us/87/8764/a876464.html`

John, H.W., Mcmeekin, R.R., Burgin, R.E., Yanowitch, R.E.: Correlation Of Occurrence of Aircraft Accidents With Biorhythmic Critically And Cycle Phase In US Air Force. US Army And Civil Aviation Pilots 48(10), 976–978 (1977), `http://www.ncbi.nlm.nih.gov/pubmed`

John, H.W., Mcmeekin, R.R., Burgin, R.E., Yanowitch, R.E.: Correlation of General Aviation Accidents With the Biorhythmic Theory. Transportation Research Board of The National Academics, NTL Integrated Search, US Army And Civil Aviation Pilots 19(3), 283–293 (1977), `http://www.worldcat.org/ocle`

Kaushik, S.P., Murthy, W.S., Shrivastava, K.K.: Biorhythmic Approach for Accident Prevention in Military Aviation. Defence Inst. of Physiology and Allied Sciences, New Delhi; Project Studies AR & DB, No. 820 (1990)

Khalil, T.M., Kurucz, C.N.: The Influences of 'Biorhythm' On Accident Occurrence and Performance. Ergonomic 20(4), 387–398 (1977)

Kurucz, C.N., Charls, W.: Probability Models For Analyzing The Effect Of Biorhythms On Accidents Occurrences. Journal Of Safety & Research, 150–155 (1977)

Michaela, A., Moffect, D.: Aviation Week and Space Technology, pp. 58–60 (August 17, 2000)

Clarke, M. (ed.): Brassey's Defence Yearbook, 1995. Centre for Defence Studies, Publisher, Brassey's (U.K.) Ltd (1995), `http://www.alibris.co.uk/search/books/` ISBN-13: 9781857531312 ISBN: 1857531310

Sacher, D.: The Influence of Biorhythmic Critically on Aircraft Mishaps. Defense Technical Information Center, Naval Postgraduate School, Monterey Calif., AD0783817 (September 1974), `http://www.oai.dtic.mil/oai`

Saket, R.K., Kaushik, (Wg. Cdr.) S.P.: Use of Gaussian Distribution Method for Biorhythmic Analysis to Prevent Aviation Accident. In: Proceedings of The 14th ISME International Conference on Mechanical Engineering in Knowledge age, Delhi College of Engineering, Delhi (India), pp. 139–147 (December 2005), `http://www.dce.ac.in`

Saket, R.K., Kaushik, (Wg. Cdr.) S.P., Singh, G.: Probabilistic Approach for Biorhythmic Analysis to Prevent Aviation Accidents. International Journal of Intelligent Defence Support Systems 1(4), 271–289 (2008), `http://www.inderscience.com/ijidss`

Shaffer, J.W., Zlotowitz, H.L., Fisher, R.S.: Biorhythm and Highway Crashes. Achieves of General Psychiatry, 41–46 (1978)

Shappell, S.A., Detwiler, C.A., Holcomb, K.A., Hackworth, C.A., Boquet, A.J., Wiegmann, D.A.: Human Error And Commercial Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using HFACS. Human Factors Engineering and man Machine System, Federal Aviation Administration, Washington DC Office Of Aviation Medicine, A568364 (July 2006), `http://www.stormingmedia.us/56/5683/A568364.html`

Thommen, G.S.: Biorhythms: Is This Your Day (1973), PHB CROWN 1987, `http://www.cycletourist.com/biochart` ISBN: 0-517-00742-8

Thompson, S.G., Ghanea-Hercock, R. (eds.): DAMAS 2005. LNCS (LNAI), vol. 3890. Springer, Heidelberg (2006) ISBN: 978-3-540-32832-2

Zimmerman, R.O.: Safety Intervention Effectiveness. DOE Scientific and Technical Information, USDOE Office of Environmental Management, UK, OSTI ID: 807467 (October 2001)

Nomenclatures:

P_F = Failure Probability

P_s = Success Probability

Pa = Performance Ability of Pilot to Fly

Pd = Probability of Demand

$\overline{Pd}$ = Probability of Mean Demand

C = Capability of the Pilot

$\overline{C}$ = Mean Capability of the Pilot

β = Constant independent of x′

φ (β) = Area under Normal Distribution Curve

$\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_i$ = Frequency of Accident Occurrence

P_{di} = i^{th} Demand Levels

$\frac{\overline{C}}{P_{dmax}}$ = Safety Factor

P_{dmax} = Peak Performance Demand

σ_C = Standard Deviation of Ability of the Pilot

σd = Standard Deviation of Demand on the Pilot

Pac = Biorhythmic Accident Probability

f_s = Safety Distribution Function

S = Safety Factor

$\overline{S}$ = Mean Safety Factor

σ_s = Standard Deviation for Safety Factor

Chapter 8
Reliability Evaluation of Defence Support Systems

R.K. Saket

Associate Professor
Department of Electrical Engineering
Institute of Technology, Banaras Hindu University,
Varanasi - 221005 (Uttar Pradesh), India

Abstract. The motivation for the preparation of this chapter is my wish to create an integrated introductory resource for the study of the systems reliability principles based on Defence Supports Systems (DSS). The focus across the chapter created is the support of design activities that lead to the production of dependable and efficient Defence equipments. Notwithstanding the emphasis upon initial study, the chapter has also been served well as a resource for practicing engineers, academicians and technocrats. Engineers and Scientists of the DSS who are involved in the design process and verification of product quality will find coherent explanations of the reliability and maintenance issues that will influence the success of the devices they design for composite DSS. This chapter is intended to provide a fundamental course in reliability theory to evaluate systems reliability in fields of the engineering and technology. It begins with an introduction to the key statistical concepts required for the implementation of a reliability analysis: both the analytical and the numerical methods used have been described with worked examples. Examples on applications of the methods are given to illustrate the advantages and limitations of the different techniques, together with case studies drawn from the author's experience of academia and consultancy. Comprehensive coverage of the basic concepts of probability theory, Intelligent Defence support system (IDSS) structures with reliability evaluations, Hazard model for failure analysis and various probability distributions of IDSS, solved as well as unsolved numerical examples based on IDSS in each sub-section have been described in this chapter.

1 Introduction

RELIABILITY is an old concept and new discipline of engineering and technology. Peoples have been called reliable if they had lived up to certain expectations, and unreliable otherwise. A reliable person would never fail to deliver what he had promised. Reliability of the system is defined through the mathematical concept of probability. Reliability is the probability of the system or device performing its function adequately for the period of time intended under the operating conditions intended [1]. According to basic definition of the systems reliability, specific

L.C. Jain et al. (Eds.): Innovations in Defence Support Systems - 2, SCI 338, pp. 241–286.
springerlink.com

attributes of the reliability are: probability of the system or device, adequate performance, period of time and operating conditions. Reliability theory as an extension of probability theory was first applied in the electronics, nuclear, space and Defence industries after world war-II, where, high reliability was a requirement from these increasingly complex systems [2]. Reliability evaluation of the IDSS has been started after the bomb attacks by America at Hiroshima and Nagasaki cities of the Japan in 1945 during world war-II. Occasionally, we observe a catastrophic failure. Fatigue failures of the fuselage of aircraft, the loss of an engine by a commercial jet, the Three Mile Island and Chernobyl nuclear reactor accidents, and the Challenger and Discovery space shuttle accidents are all widely known examples of catastrophic equipment failures [3]. The accident at Three Mile Island was precipitated by the failure of a physical component of the equipment. The relay circuit failure at the Ohio power plant that precipitated the August 2003 power blackout in the northeast United States and in eastern Canada is an example of a system failure that directly affected millions of peoples.

Reliability engineering is the study of the longevity and the failure of equipment. Principle of science and mathematics are applied to the investigation of how device age and fail. The intent is that a better understanding of device failure will aid in identifying ways in which product designs can be improved to increase life length and limit the adverse consequences of failure [4]. The key point here is that the focus is upon design. New product and system designs must be shown to be safe and reliable prior to their fabrication and use. Nowadays reliability studies are performed in almost all engineering branches. Such studies evolve applications for both repairable and non-repairable systems in all areas. Electric power systems are prime examples of system where a very high degree of reliability is expected [4]-[5]. The reliability is usually divided into adequacy and security. Adequacy relates to the existence of sufficient facilities within the system to satisfy the customer load demand. These include the facilities necessary to generate sufficient energy and the associated transmission and distribution facilities required to transport the energy to the actual customer load points. Adequacy assessment usually required probabilistic models for different parts of the power system, such as load and generation. Security relates to the ability of the system to respond to disturbances arising within that system [4]-[5].

Several authors have defined analytical approaches to modeling the effects of humans and of software on system reliability. The motivation for doing this is the view that humans cause more system failures than does equipment [6]-[7]. Nevertheless, implementation of the existing mathematical models of human and software reliability requires the acceptance of the view that probability models appropriately represent dispersion in human behavior. Probabilistic approaches for biorhythmic analysis to prevent aviation accidents have been presented in [8]-[9]. A new methodology has been developed using Gaussian distribution approach for reliability evaluation of the Defence system keeping in mind the biorhythmic effects of the pilot. The normally distributed data of the US air force were tested and analyzed based on the performance ability of the pilot and the peak demand of the performance [10]. The safety factor concept including peak demand of the pilot has presented in papers. The stepped incident-duration curve has also been utilized to evaluate the pilot's reliability on the aviation system using Simpson's $1/3^{rd}$ rule [8].

Defence applications are subject to some of the world's most demanding requirements for reliability, controllability, security, flexibility, and synchronization. The evolution of Defence processes towards network enabled systems and rapid deployment scenarios, is creating an urgent demand for highly adaptive and autonomous information support systems. In particular there are requirements for reduced manpower in support roles, autonomous IT infrastructures, and automated logistics and planning, all of which provide significant scope for an agent-oriented solution set. The workshop addresses the use of agent systems and agent applications applied to Defence scenarios in support of these requirements [11]. Unsurpassed coverage of all aspects of reliability engineering and management, including reliability issues in electronic components and systems, software, and mechanical devices have been described in [12]. It now includes specific information on how to design a product for reliability; it adds the concept of process and the tools of total quality control (TQC) to the reliability function. Practical information on proven industry practices to define and achieve reliability goals, as well as the traditional mathematics of reliability. Also included are basic tables for determining reliability, and standards and specifications used b the U.S. Department of Defense. Essential for all reliability engineers, product designers, quality engineers, and engineering mathematics, this edition of the world-renowned handbook will give the expertise needed to define and attain optimum reliability goals for Defence company's products. Specific engineering, management, and mathematics data need to design and manufacture more reliable electronic and mechanical devices as well as complete systems have been presented in this book [12]. Comment and analysis on the current agenda of Defence and security policies, five main sections: British Defence policy, European security, arms control, regional security and perspectives on security have been described in [13]. Reliability and Probabilistic Methods in Coastal and Hydraulic Engineering set out the methods which are increasingly being required by Government Agencies for river and sea Defence design and flood Defence system management [14]. It highlights the major concepts developed during the last two decades for describing uncertainty in the performance of flood and erosion Defence.

According to basic definition of the IDSS Reliability, specific attribute of reliability is the success probability of the combined system. Reliability of the various composite systems has been evaluated using probability theory [15]-[16], [21]. Fundamentals of the probability theory including various probability distributions, Probabilistic Models and Statistical Methods have been presented in this chapter for reliability analysis of the IDSS [17]-[20]. This chapter has organized as follows: Section: 2 describe basic probability theory considering conditional and independent events. Baye's theorem with IDSS solved numerical examples and unsolved questions have been presented in this section. Hazard model of the DSS has been illustrated in section: 3. Section: 4 presents life distributions of the DSS / components. Binomial distribution, poisons distribution, normal distribution, exponential distribution, gamma distribution, Weibull distribution have been included in this section. Basic concepts or systems reliability and mean time to failure have been described in section 5 and 6 respectively. Reliability of DSS structures including series, parallel and k-out-of-n system structures with numerical examples have been described in section 7. Reliability and MTTF of some special failure distribution functions with miscellaneous solved numerical examples of IDSS have been illustrated in section 8. Section 9 presents conclusion on reliability evaluation of the DSS.

2 Probability Theory: An Overview

According to basic definition of reliability, main attribute of the reliability is the probability of the system performing its function adequately. Evaluations of system reliability require probability calculations. We can formulate a mathematical expression for the reliability R (t) of a system S required to perform a mission under specified conditions.

$$R(t) = \text{probability (S will be operate during } [0, t])$$

If there are n exhaustive exclusive cases of which m are favorable to an event A, then the probability (p) of the happening of A is (m/n). As there are (n-m) causes in which A will not happen, therefore the chances of A not happening is:

$$q = (n-m)/n \quad \Rightarrow q = 1-(m/n) \quad \Rightarrow q = (1-p) \quad \Rightarrow (p + q) = 1 \qquad (1)$$

If an event is certain to happen then its probability is unity, while if it is certain not to happen, its probability is zero.

2.1 Conditional Probability

Let S be a sample space with the probability function P defined on it. Let A and B to be two events of S. The conditional probability of B, assuming that the event A has happened, is denoted by P (B/A) and defined as,

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) > 0. \qquad (2)$$

Similarly, $$P(A/B) = \frac{P(A \cap B)}{P(B)} \text{ provided } P(B) > 0 \qquad (3)$$

Remarks:

(i) From the definition of conditional probability for any two events A and B we have
$P(A \cap B) = P(A/B)\,P(B) = P(B/A)P(A)$
This relation is called multiplication theorem for probability.
(ii) For $P(B) > 0$, $P(A/B) \leq P(A)$
(iii) For n events $A_1, A_2, \ldots A_n$ we have,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)\ldots$$

$$\ldots \times P(A_n/A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

This is extension of multiplication theorem for probability.

(iv) For any three events A, B and C

$$P(A \cup B/C) = P(A/C) + P(B/C) - P(A \cap B/C)$$

Solved Numerical Examples

Problem 2.1.1. In a shooting test, the probabilities of hitting a target are $\frac{2}{5}$ for A, $\frac{1}{2}$ for B and $\frac{7}{10}$ for C. If all them fire at the same target, calculate the probability that atleast one of them hits the target.

Solution: Given $P(A) = \frac{2}{5}$; $P(B) = \frac{1}{2}$; $P(C) = \frac{7}{10}$

Then $P(\overline{A}) = \frac{3}{5}$; $P(\overline{B}) = \frac{1}{2}$; $P(\overline{C}) = \frac{3}{10}$

∴ P(at least one of them hits)

$$= P(A \cup B \cup C) = 1 - \left(\overline{A \cup B \cup C}\right)$$

$$= 1 - \left[P(\overline{A}).P(\overline{B}).P(\overline{C})\right] \quad \text{(since trials are independent)}$$

$$= 1 - \left[\frac{3}{5} \times \frac{1}{2} \times \frac{3}{10}\right] = 1 - \frac{9}{100} = \frac{91}{100} \text{ Ans.}$$

Problem 2.1.2. A manufacturer of air plane parts knows that the probability is 0.8 that on order will be ready for shipment on time, and it is 0.7 that an order will be ready for shipment and will be delivered on time. What is the probability that such an order will be delivered on time given that it was also ready for shipment on time?

Solution: Let A = event that order is ready for shipment on time.

B = event that order is delivered on time.

Here P(A) = 0.8 ; P(A $\cap$ B) = 0.7

∴ P [an order will be delivered on time given that it was also ready for shipment on time]

$$= P\left(\frac{B}{A}\right) = \frac{P(A \cap B)}{P(A)} = \frac{0.7}{0.8} = \frac{7}{8} \text{ Ans.}$$

2.2 Independent Events

Let A and B be any two events. If A is said to be independent (or statistically independent) of B, then the conditional probability of A given B is equal to the unconditional probability of A.

i.e., $$P(A/B) = P(A) \tag{4}$$

Remarks:

(i) If A and B are any to events such that $P(A) \neq 0$ and $P(B) \neq 0$ and if A is independent of B then B is independent of A.
(ii) If A and B are any two independent events than $P(A \cap B) = P(A) . P(B)$.
(iii) Conditions for Mutual Independence of n events: If $A_1, A_2, \ldots. A_n$ are events, then for their mutual independence, we should have,
(a) $P(A_i \cap A_j) = P(A_i)P(A_j), (i \neq j, i, j = 1, 2, \ldots n)$.
(b) $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2) \quad P(A_n)$

2.3 Baye's Theorem

Statement: Let $\{A_i\}$ be a sequence of mutually exclusive and exhaustive events in a sample space S such that $P(A_i) > 0, \forall_i$. Let B be any event in S with $P(B) > 0$. Then,

$$P(A_i / B) = \frac{P(A_i)P(B/A_i)}{\sum_i P(A_i)P(B/A_i)} \tag{5}$$

Proof: Let A_i be an event in S and let B any event with $P(B) > 0$.
Then by definition of conditional probability

$$P(A_i / B) = \frac{P(A_i \cap B)}{P(B)}$$

$$\Rightarrow \qquad P(B)\,P(A_i / B) = P(A_i \cap B) \tag{6}$$

and $$P(B / A_i) = \frac{P(A_i \cap B)}{P(A_i)}$$

$$\Rightarrow \qquad P(A_i)P(B / A_i) = P(A_i \cap B) \tag{7}$$

From (1) and (2), $P(A_i/B)\, P(B) = P(A_i)\, P(B/A_i)$

$$\Rightarrow \qquad P(A_i / B) = \frac{P(A_i)P(B/A_i)}{P(B)} \tag{8}$$

Now we claim that, $$P(B)=\sum_i P(A_i)P(B/A_i)$$

Since the events (Ai) are mutually exclusive and exhaustive in S we have $\bigcup_i A_i = S$ and A_i's are disjoint.

$$\therefore \qquad B = B\cap S = B\left(\bigcup_i A_i\right)=\bigcup_i (B\cap A_i)$$

$$\text{This} \Rightarrow P(B)=P\left[\cup_i (B\cap A_i)\right]=\sum_i P(B\cap A_i) \quad \text{(by axiom of probability)}$$

$$=\sum_i P(A_i)P(B/A_i) \text{ By conditional probability} \qquad (9)$$

Substitute equation (9) in equation (8), we get the required result.

Solved Numerical Examples

Problem 2.3.1. One factory F_1 produces 1000 articles for Defence support system, 20 of them being defective products, second factory F_2 produces 4000 articles, 40 of them being defective and third factory F_3 produces 5000 articles 50 of them being defective. All these articles are put in one stock pile. One of them is chosen and is found to be defective. What is the probability that it is from factory F_1?

Solution: Let A_1, A_2 and A_3 denote the events that chosen article which is produced by F_1, F_2, F_3 respectively. Here the total no. of products

$$N(S) = 1000 + 4000 + 5000 = 10{,}000$$

And
$$P(A_1)=\frac{1000}{10{,}000}=\frac{1}{10}, P(A_2)=\frac{4000}{10{,}000}=\frac{2}{5}$$
$$P(A_3)=\frac{5000}{10{,}000}=\frac{1}{2}$$

Let B denotes the event that the article chosen is defective. Then

$$P(B/A_1)=\frac{20}{1000}=\frac{1}{50}$$
$$P(B/A_2)=\frac{40}{4000}=\frac{1}{100}; P(B/A_3)=\frac{50}{5000}=\frac{1}{100}$$

To find P(A_1/B) By Baye's theorem:

$$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{\sum_{i=1}^{3} P(A_i)P(B/A_i)}$$

$$= \frac{\frac{1}{10}\times\frac{1}{50}}{\left(\frac{1}{10}\times\frac{1}{50}\right)+\left(\frac{2}{5}\times\frac{1}{100}\right)+\left(\frac{1}{2}\times\frac{1}{100}\right)} = \frac{2}{11}$$

Problem 2.3.2. A missile is rejected if the design is faulty or not. The probability that the design is faulty is 0.1 and that the missile is rejected because of faulty design is 0.95 and otherwise is 0.45. If a missile is rejected, what is the probability that it is due to faulty design?

Solution: Let F1, F2 be events that design is faulty or not and A be the event that the missile is rejected.

Then $P(F_1) = 0.1$, $P(F_2) = 0.9$ and $P(A/F_1) = 0.95$, $P(A/F_2) = 0.45$

$P(F_1/A)$ = P[rejection due to faulty design]

$$= \frac{P(F_1)P(A/F_1)}{P(F_1)P(A/F_1)+P(F_2)P(A/F_2)} = \frac{0.1\times 0.95}{0.1\times 0.95 + 0.9\times 0.45} = 0.19$$

Problem 2.3.3. There are three machines producing 10,000, 20,000 and 30,000 bullets for IDSS respectively. These machines are known to produce 1%, 2% and 1% defectives. One bullet is taken at random on a days production of the three machines are found to be defective. What is the probability that this bullet came from the third machine?

Solution: Let A_1 = bullet came from the first machine.

A_2 = bullet came from the second machine

A_3 = bullet came from the third machine

And B = taken bullet is found to be defective.

Hence n(S) = 10,000 + 20,000 + 30,000 = 60,000

$$P(A_1) = \frac{10{,}000}{60{,}000} = \frac{1}{6} \quad P(A_2) = \frac{20{,}000}{60{,}000} = \frac{1}{3}\ ; \quad P(A_3) = \frac{30{,}000}{60{,}000} = \frac{1}{2}$$

$$P\left(\frac{B}{A_1}\right) = \frac{1}{100}\ ; \quad P\left(\frac{B}{A_2}\right) = \frac{2}{100}\ ; \quad P\left(\frac{B}{A_3}\right) = \frac{1}{100}$$

To find $= P\left(\frac{A_3}{B}\right)$ **By Baye's theorem:**

$$= P\left(\frac{A_3}{B}\right) = \frac{P(A_3)P\left(\frac{B}{A_3}\right)}{\sum_{i=1}^{3} P(A_i)P\left(\frac{B}{A_i}\right)}$$

$$= \frac{\left(\frac{1}{2}\times\frac{1}{100}\right)}{\left(\frac{1}{6}\times\frac{1}{100}\right)+\left(\frac{1}{3}\times\frac{2}{100}\right)+\left(\frac{1}{2}\times\frac{1}{100}\right)}$$

$$= \frac{\left(\frac{1}{200}\right)}{\left(\frac{1}{600}\right)+\left(\frac{2}{300}\right)+\left(\frac{1}{200}\right)} = \frac{1}{200}\times\frac{600}{(1+4+2)} = \frac{3}{8}$$

2.4 Unsolved Numerical Problems

Q. 2.4.1 The probability that a new airport will get an award for its design is 0.16, the probability that it will get an award for the efficient use of materials is 0.24 and the probability that it will get both awards is 0.11.

(a) What is the probability that it will get atleast on of the two awards?

(b) What is the probability that it will get only one of two awards?

[Ans. (a) 0.29, (b) 0.18]

Q. 2.4.2 In a bolt factory, machines A, B and C manufacture respectively 25%, 35% and 40% of the total. Of their output 5, 4, 2 percents are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines A, B and C?

[Ans. 25/69, 28/69, 16/69]

Q. 2.4.3 Four techniques regularly make repairs when breakdowns occur on an automated production line. A, who services 20% of the breakdowns, makes an incomplete repair 1 time in 20 ; B, who services 60% of the breakdowns, makes an incomplete repair 1 time in 10 ; C who services 15% breakdowns, makes an incomplete repair 1 time in 10 ; and D, who services 5% of breakdowns, makes an incomplete repair 1 time in 20 ; For the next problem with the production line diagnosed as being due to an initial repair that was incomplete, what is the probability that this initial repair made by A ? **[Ans. 0.114]**

Q. 2.4.4 Three machines M1, M2 and M3 produce identical items. Of their respective output 5%, 4%, and 3% of items are faulty. On a certain day, M1 has produced 25% of the total output, M2 has produced 30% and M3 the remainder. An item selected at random is found to be faulty. What are the chances that it was produced by the machine with the highest output? **[Ans. 0.355]**

3 Hazard Model of Defence Support Systems

Hazard function is rate at which surviving units fail. The hazard function is commonly used by reliability analysts to describe the failure behaviors of the device/system. The use of the hazard function started with the concept that the population of devices displays a "bathtub shaped" hazard over the lives of the members of the population [22]. The shape is intended to illustrate the view that aging in a device population proceeds through phases. Early in the lives of the devices, failure occurs at a relatively high rate. This "infant-mortality period" is often attributed to the failure of members of the population that are "weak" as a result of material flaws, manufacturing defects, or other physical anomalies. Following the "early life" or "infant mortality" period, the device population proceeds through the "function life period" during which the hazard function is relatively low and reasonably stable [23]-[25]. Finally, towards the end of the lives of the population members, survivors fail with an increasing rate as a consequence of "wear out." Hazard model / bathtub curve can be divided into following three parts.

(1) Initial period (Burn in / debugging / infant monolith period): In burn in period, a decreasing hazard function indicates continuous reduction in the chance of an imminent failure as time passes. This model is usually valid for the initial period of a component operation, which has a decreasing hazard function for all its life.
(2) Constant Hazard Model: The constant hazard model applies to components where the chance of failure in some interval remains the same all the time. In mid-life no extraordinary accumulations of failures are expected and therefore the so called "chance failure", that have a comparatively low and constant hazard rate dominate.
(3) Wear out region: An increasing hazard function essentially indicates that the components so described become more prone to failure as they age, that is for a population of the same size surviving same time t.

$$H(t) = f(t)/R(t)$$

Where, $h(t)$= hazard function, $f(t)$=probability density function, and $R(t)$=reliability function of the device or system.

4 Life Distributions of Defence Support System: An Overview

In principle, any distribution may be used to model equipment longevity. In practice, distribution functions having monotonic hazard functions seem most realistic, and within that class, there are a few that are generally thought to provide the most reasonable models of device reliability [23]. The most common choices of life distribution models are given below.

4.1 The Binomial Distribution

Consider a trial in which there are only two possible outcomes say, a success or failure. Let p be the probability of success and (1-p) = q is the failure in any one experiment. And let X be the random variable that equals the number of success in n trials [23]-[24]. Then the probability getting x success is b(x, n, p) = ${}^nC_xp^xq^{n-x}$, where x = 0, 1, 2 ….. n and combinatorial quantities ${}^nC_x = \frac{n!}{(n-x)!x!}$ are referred to as binomial coefficients. This distribution is called binomial distribution.

Corollary: The sum of the probabilities

$$\sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} {}^nC_x p^x q^{n-x} = q^n + {}^nC_1 p^1 q^{n-1} + ... + p^n = (q+p)^n = 1 \quad (10)$$

Applications of the Binomial Distribution

(i) Number of rounds fired from a gun hitting a target.
(ii) Radar detection in the Defence support systems.
(iii) Numbers of defectives in a sample from production line.
(iv) Estimation of reliability of Defence systems.

4.2 The Exponential Distribution

The most widely used distribution function for modeling reliability is the exponential distribution. It is such a popular model of device reliability because it is algebraically simple and thus tractable and it is considered representative of the functional life interval of the device life cycle [24]. The device are expected to be absolute before reaching the wear out period, so an appropriate model of device reliability is one having constant hazard.

Definition: A continuous random variable X assuming non-negative values is said to have an exponential distribution with parameter $\alpha > 0$, if its pdf is given by

$$f(x) = \begin{cases} \alpha e^{-ax}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

The distribution function {F(x)} is given by F(x) = $\int_0^x \alpha e^{-ax}\, dx.$ (11)

4.3 Poisson Distribution

Definition: The probability density function of the Poisson variants can be obtained as limiting case of the binomial probability density function when $p \to 0$

and n → ∞. Thus Poisson distribution is the distribution of a variable x with relative frequency.

$$f(x,m)=\frac{e^{-m}m^{x}}{x!}\ and\ \sum_{x=0}^{\infty}f(x,m)=\sum_{x=0}^{\infty}\frac{e^{-m}m^{x}}{x!}=e^{-m}\sum_{x=.0}^{\infty}\frac{m^{x}}{x!}=e^{-m}e^{m}=1.$$

This is an exponential probability distribution with only one parameter 'm'. (12)

Applications of Poisson Distribution:

(i) Spatial distribution of bomb hits.
(ii) Number of fragments from a shell hitting a target.
(iii) Arrival pattern of defective vehicles in a Defence workshop.
(iv) Demand pattern for certain spare parts.

Problem 4.3.1. A manufacturer of bullets knows that 5% of his product is defective. If he sells bullets in boxes of 100 and guarantees that not more than 10 bullets will be defective, what is the approximate probability that a box will fail to meet the guaranteed quality?

Solution: We are given n = 100
Let p = probability of a defective bullets = 5% = 0.05
∴ m = mean number of defective bullets in a box of 100 = np = 100 x 0.05 = 5.
Since p is small, we can use Poisson distribution. Probability of x defective bullets in a box of 100 is P(X = x) = $\frac{e-mmx}{x!}=\frac{e-55x}{x!},\ x=0,1,2,...$
Probability that a box will fail to meet the guaranteed quality is,

$$P(X>10)=1-P(X\le 10)=1-\sum_{x=0}^{10}\frac{e^{-5}5^{x}}{x!}=1-e^{-5}\sum_{x=0}^{10}\frac{e^{-5}5^{x}}{x!}=1-e^{-5}\sum_{x=0}^{\infty}\frac{5^{x}}{x!}$$

Problem 4.3.2. In a certain factory turning razor blades, there is a small chance of $\frac{1}{500}$ for any blade to be defective. The blades are in packets of 10. Use Poisson distribution to calculate the approximate umber of packets containing (i) no defective (ii) one defective (iii) 2 defective blades respectively in a consignment of 1000 packets.

Solution: Given $p=\frac{1}{500}$; $n=10,\ and\ \ N=1000$

∴ Mean m = n x p = $10\times\frac{1}{500}=\frac{1}{50}=0.02.$

And the Poisson distribution is, $p(x) = \dfrac{e^{-m} m_x}{x!} = \dfrac{e^{-0.02}(0.02)^x}{x!}$

(i) P (no defective) = p(0) = $\dfrac{e^{-0.02}(0.02)^0}{0!}$ = 0.980198.

∴ The total number of packets containing no defective blades in a consignment of 1000 packets = N x p(0) = 1000 x 0.98019 = 9802.

(ii) P(one defective) = p(1) = 0.0196 ∴ The total number of packets containing one defective blade in a consignment of 1000 packets = N x p(1) = 1000 x 0.0196 = 20.

(iii) P(two defective) = p(2) = 0.00196. ∴ The total number of packets containing two defective blades in a consignment of 1000 packets = N x p(2) = 1000 x 0.00196 = 2.

Problem 4.3.3. In a sample of large number of missile parts produced by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples how many would be expected to contain atleast 3 defectives parts?

Solution: Mean = m = 2

Here the r.v X follows Poisson distribution.

The probability function is,

$p(x) = \dfrac{e^{-2} 2^x}{x!}$, Where x is the no. of defectives

P[a part contain atleast 3 defectives]

$$= P[X = 3] + P[X = 4] + \ldots\ldots.$$

$$= 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\}$$

$$= 1 - \left\{ \frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} + \frac{e^{-2}.2^2}{2!} \right\}$$

$$= 1 - (0.135 + 0.270 + 0.270) = 0.325$$

∴ Out of 1000, the number of parts with atleast three defectives = 0.325 x 1000 = 325 Ans.

Problem 4.3.4. It is known that the probability of an item produced by a certain Defence machine will be defective is 0.05. If the produced items are sent to the market in packets of 20, find the number of packets containing atleast, exactly and almost 2 defective items in a consignment of 1000 packets using Poisson's approximation to binomial distribution.

Solution: Here n = 20; p = 0.05. Then m = np = 20 x 0.05 = 1.

Let X is number of defective items. X follows Poisson distribution.

∴ The probability function is

$$P(X = x) = \frac{e^{-m} m^x}{x!} = \frac{e^{-1}}{x!},$$

P [packet containing atleast 2 defectives]

$= P[X \geq 2] = 1 - P[X < 2]$
$= 1 - \{P[X = 0] + P[X = 1]\}$
$= 1 - \{e^{-1} + e^{-1}\} = 1 - 0.7354 = 0.2642$

∴ The number of packets containing atleast 2 defectives = 1000 x 0.2642

= 264 (approximately)

P[a packet containing exactly 2]

$$= P[X = 2] = \frac{e^{-1}}{2!} = 0.1839$$

∴ The number of packets containing exactly two defectives = 1000 x 0.1839 = 184.
$P[X \leq 2] = P[X = 0] + P[X = 1] + P[X = 2]$

$$= e^{-1} + e^{-1} + \frac{e^{-1}}{2!} = 0.9193$$

∴ The number of packets containing almost 2 defectives = 1000 x 0.9193 = 919.

4.4 Geometric Distribution

Suppose in an experiment, the number of trials, n is not fixed. Clearly, if the first success is to come on the X^{th} trial, it has to be preceded by (x - 1) failures, and if the probability of a success is p, the probability of (x - 1) failures in (x - 1) trials is (1 - p)x-1.

Definition: A random variable X is said to have a geometric distribution if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = \begin{cases} pq^{x-1}, & x = 1,2,3,...0 < p \leq 1, q = 1 - p \\ 0, & otherwise. \end{cases} \tag{13}$$

This probability distribution is called the geometric distribution.

Problem 4.4.1 Suppose that a trainee soldier shoots a target in an independent fashion. If the probability that the target is shot on any one shot is 0.7.

(i) What is the probability that the target would be hit on 10th attempt?
(ii) What is the probability that it takes him less than 4 shots?
(iii) What is the probability that it takes him an even number of shots?

Solution: Here p = 0.7 then q = 1 – p = 0.3

(i) The probability that the target would be hit on 10th attempt = $(0.7)\ (0.3)^{10-1}$ = $(0.7)\ (0.3)^9 = 0.000014$.
(ii) The probability that it takes him less than 4 shots is

$$P(X < 4) = \sum_{n=1}^{4-1} pq^{n-1} = \sum_{n=1}^{3} (0.7)(0.3)^{n-1}$$
$$= (0.7)[(0.3)^0 + (0.3)^1 + (0.3)^2] = (0.7)[1 + 0.3 + 0.09]$$
$$= (0.7)(1.39) = 0.973.$$

(iii) The probability that it taken him an even number of shots.

$$\sum_{n=1}^{\infty} (0.7)(0.3)^{2^{n-1}} = (0.7)(0.3)\sum_{n=1}^{\infty} (0.3)^{2^{n-2}} = (0.7)(0.3)\frac{1}{1-(0.3)^2} = 0.23$$

4.5 The Weibull Distribution

The Weibull distribution is closely related to the exponential distribution. It has been found that the distribution provides a reasonable model for the life length of very many devices. The Weibull distribution is very widely used in reliability modeling [22]-[25]. It has the advantages of flexibility in modeling various types of hazard behavior and of algebraic tractability. Weibull distribution is one possible realization of the extreme value distribution.

Definition: The probability density function of Weibull distribution of a random variable X is given by $f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}, & for\, x > 0, \alpha > 0, \beta > 0 \\ 0, & elsewhere \end{cases}$ (13)

To demonstrate this relationship, we evaluate the probability that a random variable having the Weibull distribution will take on a value less than a namely the integral $\int_0^a \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}\ dx.$

Making the change of variably y = x^{β} we get $\int_0^{\alpha^{\beta}} \alpha e^{-\alpha y} dy = 1 - e^{-\alpha a^{\beta}}$ and it can be seen that y is value of a random variable having an exponential distribution.

Mean and Variance of Weibull distribution:

The mean of the Weibull distribution having the parameters α and β may be obtained by the evaluating the integral.

$$Mean\ \mu = \int_0^x x\left(\alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}\right)dx \quad = \quad \int_0^{\infty} \alpha\beta x^{\beta}\ \ e^{-\alpha x^{\beta}}$$

Put $u = \alpha x^{\beta}$ $\Rightarrow$ $du = \alpha\beta x^{\beta-1}\,dx$

Also if $x = 0$ $\Rightarrow$ $u = 0$ and $x = \infty$ $\Rightarrow$ $u = \infty$

$$= \frac{\alpha\beta}{\alpha^{1/\beta}}\int_0^{\infty}\left(\frac{u}{\alpha}\right)e^{-u}\,\frac{1}{\beta}u^{(1/\beta)-1}du$$

$$=\alpha^{-1/\beta}\int_0^{\infty}u^{1/\beta}\,e^{-u}du = \alpha^{-1/\beta}\int_0^{\infty}e^{-u}u^{\left(\frac{1}{\beta}+1\right)-1}\,du$$

$$=\alpha^{-1/\beta}\Gamma\left(1+\frac{1}{\beta}\right). \tag{14}$$

$$\left(\because\ \Gamma n = \int_0^{\infty}e^{-x}x^{n-1}dx\right)$$

To find variance:

Now $E(X^2) = \int_0^{\infty}\alpha\beta x^{\beta+1}e^{-\alpha x^{\beta}}\,dx$ to evaluate use the substitution $u = \alpha x^{\beta}$

$\alpha^{-2/\beta}\Gamma\left(1+\frac{2}{\beta}\right)$

Then, variance

$$\sigma^2 = E\left[(X-\mu)^2\right] = E\left[X^2\right]-\mu^2 = \alpha^{-2/\beta}\left[\Gamma\left(1+\frac{2}{\beta}\right)-\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2\right] \tag{15}$$

4.6 The Normal Distribution

Definition: A continuous variable X is said to follow a normal distribution or Gaussian distribution with parameters μ and σ, if its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad -\infty < x < \infty \tag{16}$$

Symbolically, X follows N (μ,σ), Here mean of X = μ and S.D = σ

The total area bounded by the normal curve and the X-axis is 1. Hence the area under the curve between two ordinates X = a, and X = b, where a < b represents the probability that X lies between a and b. This probability is denoted by P[a < X < b].

Suppose the variable X is expressed in terms of standard units [$Z = (X - \mu)/\sigma$], equation (16) is replaced as $\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, -\infty < Z < \infty$ and is called standard normal distribution. This is obtained by putting $\mu = 0$ and $\sigma = 1$ and by changing x into Z, i.e., if X has the distribution.

N (μ,σ) then $Z = \left(\frac{X-\mu}{\sigma}\right)$ has the distribution N (0, 1).

Some properties of the normal distribution given by equation (1) are listed below:

$$\text{Mean} = \mu,\ \text{Variance} = \sigma^2,\ \text{S.D.} = \sigma$$

Mean deviation about the mean = $\int_{-\infty}^{\infty} |x-\mu| f(x)dx = \sqrt{\frac{2}{\pi}}\sigma = \frac{4}{5}\sigma \quad (approx.)$

Quartile deviation = $\frac{2}{3}\sigma$

Moment generating function of N (μ, σ)

$$M_x(t) = M_{\sigma Z+\mu}(t) \qquad \left(\because Z = \frac{X-\mu}{\sigma}\right)$$

$$= e^{\mu t} M_z(\sigma t)$$

$$= e^{\mu t} . e^{\sigma 2t2/2} \qquad \text{(since Mz(t) = } e^{t^2/2} \text{ verify it)}$$

$$= e^{t\left(\mu+\frac{\sigma^2 t}{2}\right)} = 1 + \frac{t}{1!}\left(\mu+\frac{\sigma^2 t}{2}\right) + \frac{t^2}{2!}\left(\mu+\frac{\sigma^2 t}{2}\right)^2 + \ldots\ldots$$

$\therefore E(X) = \mu$; $E(X^2) = \sigma^2 + \mu 2, \ldots$

Thus we get $\mu_1 = 0$, $\mu_2 = \sigma^2$, $\mu_3 = 0$; $\mu_4 = 3\sigma^4$ etc.

Applications of normal distribution

(i) Computation of hit probability of a shot.
(ii) Statistical inference in almost every branch of science and Technology.
(iii) Calculation of errors made by chance in experimental measurements.
(iv) Reliability evaluation of composite systems.

Solved Examples:

Problem 4.6.1. The break down voltage X of a randomly chosen diode of a particular type and use in the satellites is known to be normally distributed with μ = 40 volts and σ = 1.5 volts. What is the probability that (i) the break down voltage will be between 39 and 42 volts (ii) between 40 and 43 volts.

Solution: Here X follows N (40, 1.5).

(i) To find P[39 ≤ X ≤ 42]
First we express the event 39 ≤ X ≤ 42 in equivalent form by standardizing:

$$39 \le X \le 42 \text{ becomes } \frac{39-40}{1.5} \le \frac{X-40}{1.5} \le \frac{42-40}{1.5}$$

i.e. $-0.67 \le Z \le 1.33$ where $Z = \dfrac{X-40}{1.5}$

$\therefore$ P (39 ≤ X ≤ 42) = P(-0.67 ≤ Z ≤ 1.33)
= ϕ(Z = 1.33) - ϕ(Z = -0.67)
= 0.9082 – 0.2514 = 0.6568

(ii) $P(40 \le X \le 43) = P\left[\frac{40-40}{1.5} \le \frac{X-40}{1.5} \le \frac{43-40}{1.5}\right]$

= ϕ[0 ≤ Z ≤ 2] = ϕ [Z = 2] - ϕ[Z = 0]
= 0.4773 – 0 = 0.4773.

Problem 4.6.2. The life time of certain kinds of electronics devices have a mean of 300 hours and a S.D. of 25 hours assuming that the distribution of these life times which are measured to the nearest hour can be approximated closely with a normal curve.

Find the probability that any one of the electronic devices will have a life time of more than 35 hours.

(i) What percentage will have life times of 300 hours or less?
(ii) What percentage will have life times from 220 or 260 hours?

Solution: Here X follows N (μ = 300, σ = 25)

(i) $P(X > 350) = \left[\frac{X-300}{25} > \frac{350-300}{25}\right]$

= ϕ[Z > 2] = 1 - ϕ(Z ≤ 2)
= 1 -0.9772 = 0.0228

$P(X = 300) = \left[\frac{X-300}{25} > \frac{300-300}{25}\right]$

= ϕ(Z = 0) = 0.5000.

$\therefore$ The required percentage = 0.5 x 100 = 50%.

(ii) P(220 ≤ X ≤ 260) = ϕ (-3.2 ≤ Z ≤ -1.6)
= ϕ(-1.6) - ϕ (-3.2)
= [1 - ϕ(1.6)] – [1 - ϕ(3.2)]
= [1 – 0.9452] – [1 – 0.9903]
= 0.0548 – 0.0007 = 0.0541

$\therefore$ The required percentage = 0.0541 x 100 = 5.41%

4.7 The Gamma Distribution

Definition: Continuous random variables X which is distributed accordingly to the probability law.

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{\alpha-1}}{\Gamma(\alpha)}; & \alpha > 0, 0 < x < \infty \\ 0, & otherwise \end{cases} \tag{17}$$

is known as a Gamma variable with parameter α and referred to as Γ(α) variants and its distribution is called the Gamma distribution.

Remarks:

The function f(x) defined above represents a probability function as given below.

$$\int_0^{\infty} f(x)dx = \frac{1}{\Gamma(\alpha)} \quad \int_0^{\infty} e^{-x} x^{\alpha-1} dx = 1 \ .$$

A continuous random variable having the following pdf is said to have a gamma distribution with parameterα, u

$$f(x) = \begin{cases} \dfrac{\alpha^u}{\Gamma(u)} e^{-ax} x^{u-1}; & \alpha, u > 0, 0 < x < \infty \\ 0, & otherwise \end{cases} \tag{18}$$

5 Basic Concepts of System Reliability

As we know that main attributes of the reliability are: Probability, adequate function, period of time and operating condition of the system. The reliability of a system or a component will often depend on the length of time it has been in service. Thus, of fundamental importance in reliability studies are the failure-time distribution that is the distribution of the time to failure of a component under given environmental conditions. The component may be repairable or non-repairable. A component will be termed repairable if it can be restored to its original condition after it has failed, without affecting system operation [26]-[27]. Availability, maintainability, mean time to failure (MTTF), mean time to repair (MTTR), mean up time (MUT), mean down time (MDT), mean time between failure (MTBF) and loss of load probability (LOLP) are the main key words of the reliability[28]-[30]. A useful way to characterize this distribution is by means of its associated instantaneous failure rate. To develop this concept, let a component be put into operation at some specified time, say t = 0, and let $f(t)$ be the probability density of the time to failure of a given component that is the probability that the component will fail between times to failure of a given component, that is the probability that the component will fail between times t and t+Δt is given by f(t) Δt. Then the probability that the component will fail on the inter from 0 to t is given by the cumulative distribution function of $t, F(t) = \int_0^t f(x)dx$ and the reliability function, expressing the probability that it survives to time t, is given by

$$R(t) = 1\text{-}F(t) \tag{19}$$

The probability of success, Thus the probability that the component will fail in the interval fro t to t+ Δt is F(t+Δt) – F(t), and conditional probability of failure in this interval, given that the component survived to time t, is expressed by

$$\frac{F(t+\Delta t)-F(t)}{R(t)}.$$

Dividing by Δt , we find that the average rate of failure in the interval from t to t+ Δt , given that the component survived to time t is

$$\frac{F(t+\Delta t)-F(t)}{\Delta t}\cdot\frac{1}{R(t)}$$

Taking the limit as $\Delta t \longrightarrow 0,$ we then get the instantaneous failure rate, or simply the failure rate

$$Z(t)=\frac{F^{'}(t)}{R(t)}=\frac{1}{R(t)}\frac{dF}{dt}(t)\cdot \tag{20}$$

Finally, observing that $f(t)=\dfrac{dF(t)}{dt}$, we get the relation

$$Z(t)=\frac{f(t)}{1-F(t)} \tag{21}$$

(From (1) and (2)), which expresses the failure in terms of the failure time distribution. Equation (3) is general equation for failure rate function.

Let us now derive an important relationship expressing the failure – time density in terms of the failure rate function. Making use of the fact that R (t) = 1 – F(t) and hence, that F'(t) = - R$^{'}$(t), (2) can be written as

$$Z(t)=-\frac{R^{'}(t)}{R(t)} \tag{22}$$

Solving this differential equation for R(t), we obtain,

$$\log R(t) =- \int_0^t Z(x)ax$$

$$\Rightarrow \qquad R(t)= e^{\int_0^t Z(x)ax} \tag{23}$$

And, making use of the relation f(t) = Z(t) R(t),

We finally get f(t) = Z(t) exp $[-\int_0^t Z(x)ax]$

Equation (6) is called general equation for failure – time distribution.

6 Mean Time to Failure (MTTF) of Defence Components

We are often interested in knowing the mean time to failure (MTTF) of component rather that the complete failure details. This parameter will be assumed to be the same for all the components which are identical in the the design and operate under identical conditions. If we have life tests information on a population of N items with failure time t_1, t_2.... T_n, then the MTTF is defined as:

$$MTTF = \frac{1}{N}\sum_{i=1}^{n} t_i \tag{24}$$

However, if a component is described by its reliability function, then the MTTF is given by the mathematical expectation of the random variable T describing the time to failure of the component. Therefore,

$$MTTF = E[T] = \int_0^\infty t\, f(t)dt \tag{25}$$

But $f(t)\frac{dF(t)}{dt} = -\frac{dR(t)}{dt} \cdot Hence$

$$MTTF = -\int_0^\infty t\, dR(t)dt = -[tR(t)]_0^\infty + \int_0^\infty R(t)dt$$

$$= \int_0^\infty R(t)dt \qquad (\because R(\infty) = 0)$$

$$MTTF = \int_0^\infty R(t)dt \tag{26}$$

The MTTF can also be computed using the Laplace transform of R(t),

i.e. $MTTF = \int_0^\infty R(t)dt \underset{t\to\infty}{Lt} \int_0^\infty R(t)dt$

However, $\underset{t\to\infty}{Lt} \int_0^t R(x)dx = \underset{s\to 0}{LT} R(s)$,

Thus, MTTF = $\underset{s\to 0}{Lt} R(s)$ (27)

Remark: Var (T) = E [T – E (T)]^2 = E[T^2] – [E(T)]^2 $\int_0^t t^2 f(t)dt - \quad (MTTF)2$

Unsolved Problem:

Problem 6.1. The time to failure in operating hours of a critical solid-state power unit of the missile has the hazard rate function $\lambda(t) = 0.003\left(\frac{t}{500}\right)^{0.5}$, *for* $t \geq 0.$

(i) What is the reliability if the power unit must operate continuously for 50 hours? [Ans: R(50)=0.9689]

(ii) Determine the design life is a reliability of 0.90 is desired. [Ans: t=111.54 hrs]

(iii) Compute MTTF. [Ans: 451.65 hrs]

(iv) Given that the unit has operated for 50 hours, what is the probability that it will survive a second 50 hours of operation? [Ans: P=0.9439]

7 Reliability of Defense Support System Structures

Among reliability specialists, it is generally accepted that there are four generic types of structural relationship between a device and its components. These are Series, parallel, series-parallel and k out of n systems. In this section, we shall discuss the system reliability in respect of these simple but relatively important cases. All models are based on the assumptions that the components fall independently of each other, i.e. the failure of one component does not change the failure of other components.

7.1 Series Systems

The simplest and most commonly encountered configuration of component is the series system. The formal definition of a series system is: A series system is one in which all components must function properly in order for the system to function properly. Series or non-redundant system is one in which the components of the system are connected. Consider a system having a total of n components. Such systems are represented as shown in the following figure for the purpose of reliability estimation.

In series configuration, all the components must function for the system to function. In other words the failure of any component causes system failure. Many complex systems can be reduced to such simple structure.

Let E_i denotes the event that the component i is good (i.e. functions satisfactorily) and E_i the event that the component i is bad. The event representing system success is then the intersection of E_1, E_2, ….. E_n. Let $R_i(t)$ be the reliability of the

i^{th} component in the series, i.e. $R_i(t) = P_r(E_i)$. Then the reliability of the system is the probability of this event is given by

$R_s = P_r(E_1 \cap E_2 \cap \dots\dots\dots \cap E_n)$
$P_r(E_1)P_r(E_2)\dots\dots P_r(E_n)$ (since components are independent)

This can be evaluated using failure events E_i also. In this case

$R_s = 1 -$ (probability of the system failure)

The system fails if any of the components fail and therefore

$R_s = 1 - P_r\,(E_1 \cup E_2 \cup \dots\dots\dots \cup E_n)$

The time-dependent reliability function is

$R_s(t) = p_1(t)p_2(t)\dots\dots\dots p_n(t)$,

Where $p_i(t)$ is the probability that the component i is good at time t.

If time to failure of components are exponentially distributed, then in this case

$p_i(t) = e^{-\lambda_i t}$

And $$Rs(t) = e^{-\lambda_1 t} . e^{-\lambda_2 t} \dots\dots\dots\dots . e^{-\lambda_n t} = e^{-t\sum_{i=1}^{n}\lambda_i} \tag{28}$$

Mean time to failure of the system is

$$MTTF = \int_0^{\infty} R_s(t)dt = \int_0^{\infty} e^{-t\sum_{i=1}^{n}\lambda_i} dt = \frac{1}{\sum_{i=1}^{n}\lambda_i} = \frac{1}{\sum_{i=1}^{n}\frac{1}{T_i}} \tag{29}$$

Where T_i is the mean life of the component i.

For any general hazard model

$$R(t) = \exp\left[-\int_0^t \sum_{i=1}^{n} Z_i(x)dx\right] \tag{30}$$

In most series systems, the components are independent with respect to their probabilities of proper function. The system reliability function is an increasing function of the component reliability values and is a decreasing function of the number of components.

Solved Numerical Problems:

Problem 7.1.1. A composite system consists of five independent components in series, each having a reliability of 0.970. What is the reliability of the system?

Solution: The reliability of the five components series system is

R (t) = (0.970) x (0.970) x (0.970) x (0.970) x (0.970) = $(0.970)^5$ = 0.859 Ans.

Problem 7.1.2. A component has 99% reliability with constant failure rate. Determine the maximum number of component that can be connected in series to maintain 95% system reliability?

Solution: $(0.99)^n$ 0.95. => n log (0.99)= log (0.95) => n=5 Ans.

7.2 Parallel Systems

The second type of structure is the parallel structure. The conceptual analog is the electrical circuit, and the definition is: A parallel system is one in which the proper function of any one component implies system function. One example of a parallel system is the set of two engines on a two-engine airplane. As long as at least one engine functions, flight is sustained. However, this example implies that simply maintaining flight corresponds to proper function. Another example that is more appealing is the fact that the communications satellites presently in use have triple redundancy for each communication channels. That is, there copies of each set of transmitting components are installed in the satellite and arranged in parallel in order to assure continued operation of the channel. It is appropriate to mention the fact that the parallel arrangement of components is often referred to as redundancy. This is because the proper function of any of the parallel components implies proper function of the structure. Thus, the additional components are redundant until a component fails. Often but not always, the parallel components are identical. A distinction is made between redundancy obtained using a parallel structure in which all components function simultaneously and that obtained using parallel components of which one functions and the others wait as standby units until the failure of functioning unit. In parallel configurations of the Defence support systems,

1. If all the components have the same reliability, them

$$R(t) = 1 - [1 - p(t)]^n$$

2. In the case of constant failure rates

$$R(t) = 1 - [1 - e^{-\lambda t}]^m$$

3. The mean time to failure of the system is

$$\text{MTTF} = \int_0^\infty \{1 - \left[1 - e^{-\lambda t}\right]^n\} dt$$

Putting $(1 - e^{-\lambda t}) = x$, we get = $\dfrac{1}{\lambda}\int_0^1 \left[\dfrac{1 - x^n}{1 - x}\right] dx$

$$= \frac{1}{\lambda}\int_0^1 \left(1 + x + x^2 + \ldots\ldots + x^{n-1}\right)dx$$

$$= \frac{1}{\lambda}\int_0^1 \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots\ldots\ldots + \frac{1}{n}\right) = \frac{1}{\lambda}\sum_{i=1}^{n}\frac{1}{i}$$

When the unit reliabilities are unequal,

$$MTTF = \int_0^\infty \left[1 - \left\{\left(1 - e^{-\lambda_1 t}\right)\left(1 - e^{-\lambda_2 t}\right)\ldots\ldots\left(1 - e^{-\lambda_n t}\right)\right\}\right]dt$$

Simplifying, we get

$$= \sum_i \frac{1}{\lambda_i} - \sum_{i<j}\frac{1}{\lambda_i + \lambda_j} + \sum_{i<j<k}\frac{1}{\lambda_i + \lambda_j + \lambda_k} - \ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots + \left(-1\right)^{m-1}\frac{1}{\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n} \qquad (31)$$

The system reliability function for a parallel system is increasing in both the component reliability values and in the number of components.

Solved Problem:

Problem 7.2.1. A component has 95% reliability for a period of operation. Determine the minimum number of components connected in the parallel so that the combination reliability remains at least 99%?

Solution: Rp= [1- (1-r)^n]. => 0.99 = [1- (1- 0.95)^n]. => n=2 Ans.

Unsolved Problem:

Problem 7.2.1. A bulk IDSS system consisting of several identical components connected in parallel is to have a failure rate of at most 4 x 10^{-4} per hour. What is the least number of components that must be used if each has a constant failure rate of 9 x 10^{-4}. [Ans.: n=5 and the failure rate is 3.94x10^{-4}]

7.3 Series-Parallel (k-out-of-n) Systems

The third type of system structure is the k-out-of-n structure. There is no obvious conceptual analog for this structure. A formal definition of it is: A k-out-of-n system is one in which the proper function of any k of the n components that comprise the system implies proper system function. A system in which k-subsystems are connected in parallel where each subsystem has n-components connected in series. An example of a k-out-of-n system is the rear axle of a large tractor-trailer on which the functioning of any three out of the four wheels is sufficient to assure mobility. Another example is the fact that some (1-k) electronic memory arrays are configured so that the operation of any 126 of the 128 memory address corresponds to satisfactory operation.

Assuming the resistibility of the subsystem $S_i = P_i$ and the reliability of the path $i = R_i$,

$$R_1 = p_1p_2;\ R_2 = p_3p_4;\ R_3 = p_5p_6$$

Then the resistibility of the entire structure is: $R= [1-(1- R_1)\ (1- R_2)\ (1-R_3)\ldots]$.

In general, $R = [1 - (1 - R_1)\ (1 - R_2)\ (1- R_3)\ldots\ldots\ldots\ (1 - R_k)]$.

Remark: (i) A group contains n components in parallel, then $Rp= [1 - (1 - r)^n\ ]$. If N such groups are connected in series, then $Rp = [1 - (1 - r)^n\ ]^N$ … 32 (a)

Remark: (ii) A group contains n components in series, then $Rs= [1 - (1 - r)^n\ ]$. If N such groups are connected in parallel, then $Rs = [1 - \{(1 - r)^n\}^N]$ …32 (b).

Unsolved Problem:

Problem 7.3.1. Five components are connected in series in the Indian Missile System. N such groups are connected in parallel for reliability improvements of the system. Identical components have 95% reliability. Determine the minimum number (N) of series groups connected in parallel so that combinations reliability remains at least 99.99%. (Ans. 11 groups)

8 Failure Distribution Functions and Reliability of IDSS

In principle, any distribution function may be used to model equipment longevity. In practice, distribution functions having monotonic hazard functions seem most realistic, and within that class, there are a few that are generally thought to provide the most reasonable models of device reliability. The most common choices of life distribution models have been described as follows [3]:

8.1 Reliability of Exponential Distribution Systems

The most widely used distribution function for modeling reliability is the exponential distribution. It is algebraically simple and thus tractable, and is considered representative of the functional life interval of the device life cycle. If we make the exponential assumption about the distribution of failure times, some very useful result can be derived connecting the Mean Time between the Failures (MTBF), the mean time between the failure of series and parallel systems. We shall first have to obtain a relation expressing the reliability of a component in terms of its service time T. Making use of the fact that

$$R(t)=1-F(t)=1-\int_0^t f(x)\,dx \qquad \left(\because f(t)=\frac{dF(t)}{dt}\right)$$

If the time to failure T follows an exponential distribution with parameterα, then its PDF is given by f(t) - $\alpha e^{\alpha t}$, t≥0.

Then, $$R(t) = 1 - \int_0^t \alpha e^{-\alpha x}\,dx = 1 - \alpha\left[\frac{e^{-ax}}{-\alpha}\right]_o^t = 1 + \left[e^{-\alpha t} - 1\right] = e^{-\alpha t} \quad (33\text{-a})$$

Conversely, when $$Z(t) = \frac{f(t)}{R(t)} = \frac{\alpha e^{\alpha t}}{e^{-\alpha t}} = \alpha, cons\tan t,$$

We get, from (6), $f(t) = \alpha e^{\int_0^t \alpha dx}$

$$\Rightarrow \qquad f(t) = \alpha e^{-\alpha t}, t \geq 0 \cdot \qquad (33\text{-b})$$

Due to this property, the exponential distribution is often referred to as constant failure rate distribution in reliability contexts.

Therefore $$MTTF = E(T) = \frac{1}{\alpha} \qquad (34\text{-a})$$

And $$\operatorname{var}(T) = \sigma_T{}^2 = \frac{1}{\alpha^2}$$

Also $$R(t/T_0) = \frac{R(T_0 + t)}{R(T_0)} = \frac{e^{-\alpha(T_0+t)}}{e^{-\alpha T_0}} = e^{\alpha t} \cdot \qquad (34\text{-b})$$

This means that the time to failure of component is not dependent on how long the component has been functioning. In other words the reliability of the component for the net 1000 hours, say, is the same regardless of whether the component is brand new or has been operating for several hours. This property is known as the memory less property of the constant failure rate distribution.

Solved Numerical Problems:

Problem 8.1.1. A component has MTBF = 100 hours and MTTR = 20 hours with both failure and repair distributions exponential. Find the availability and unavailability of the component after a log time.

Solution: Given MTBF = $\frac{1}{\lambda} = 100 \quad \Rightarrow \quad \lambda = 0.01$

$$MTTR = \frac{1}{\mu} = 20 \qquad \Rightarrow \qquad \mu = 0.05$$

∴ The component availability

$$A(\infty) = \frac{MTBF}{MTBF + MTTR} = \frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + \frac{1}{\mu}} = \frac{100}{100 + 20} = 0.83$$

The component unavailability

$$\overline{A}(\infty) = \frac{\lambda}{\lambda + \mu} = \frac{0.01}{0.01 + 0.05} = \frac{0.01}{0.06} = 0.1666\ .$$

Problem 8.1.2. A mechanical pumping device has a constant failure rate of 0.023 failures per hour exponential repair time with a mean of 10 hours. If two pumps operate in an active redundant configuration, determine the system MTTF and the system reliability for days.

Solution: For the active redundant configuration

$$R(t) = \frac{m_1}{m_1 - m_2} e^{m_2 t} = \frac{m_2}{m_1 - m_2} e^{m_1 t}$$

Where $$m_1, m_2 = \frac{1}{2}\left\{-(3\lambda + \mu) \pm \sqrt{\lambda^2 + 6\lambda\mu + \mu^2}\right\}$$

Here $\lambda = 0.023$ and $\mu = \frac{1}{10} = 0.1$

$$m_1, m_2 = \frac{1}{2}\left\{-(0.069 + 0.1) \pm \sqrt{(0.023)^2 + 6 \times (0.023) \times (0.1) + (0.1)^2}\right\}$$

$$= -\,0.0065,\ -0.1625$$

Using these values in (1), we have

$$R(t) = \frac{-0.0065}{0.156} e^{-0.1625t} + \frac{0.1625}{0.156} e^{-0.0065t}$$

3 days = 72 hours.

∴ $R(72) = -0.0417 \text{ x } e^{-11.7} + 1.0417 \text{ x } e^{-0.468} = 0.6524$

$$MTTF = \frac{3\lambda + \mu}{2\lambda^2} = \frac{3 \times 0.023 + 0.1}{2 \times (0.023)^2} = 159.7 \quad hours.$$

Problem 8.1.3. An integrated circuit of intelligent system has a constant failure rate of 0.02 per thousand hours. What is the probability that it will operate satisfactorily for at least 20,000 hours? What is the 5000 hour reliability of a component consisting of 4 such chips connected in series?

Solution: Here failure rate of integrated-circuit chip is a constant.

Then failure time distribution is exponential, with parameter $\lambda = 0.02$.

And pdf is given by f(t) = 0.02 e-0.02xt, $t \geq 0$

(i) $\Pr [t \geq 20]$ $= \int_{20}^{\infty} f(t) dt$

$$= \int_{20}^{\infty} 0.02 \ e^{-0.02xt} \ dt$$

$$= -\left[e^{0.02xt}\right]_{20}^{\infty} = e^{-0.02 \times 20} = 0.6703$$

(ii) Reliability $R(t) = e^{-\lambda t}$

Here 4 chips are connected in series.

∴ $R(t = 5) = e^{-5 \text{ x } 4 \text{ x } 0.2} = 0.183$

Problem 8.1.4. A construction company is making repairs on a bridge and employing a single welder; however, to increase the reliability of the welding operation, the welder begins the job with three arc welding units, each possessing a failure rate of 0.32 per month. Assume unit failures are independent and the welder utilizes the units in a standby mode. Find the MTTF, the standard deviation of time to failure, and the probability the welder will not be without a welding unit for the duration of the job (six months). Assume that once a unit breaks down, it is not brought back to the project after repair.

Solution: Here failure rate λ = 0.32 per month.

So the time to failure T follows an exponential distribution with parameter λ then the pdf is given by f(t) = 0.32 $\times e^{-0.32t}, \geq 0$

$$And\, R(t) = e^{-\lambda t} = e^{0.32t}$$

The welder begins the job with three arc-welding units

$$MTTF = \int_{o}^{\infty} 3e^{-0.32t} dt = 9.375$$

$$Var(T) = {\sigma_T}^2 = \frac{3}{\lambda^2} = \frac{3}{(0.32)^2} = 29.297.$$

$$P(T > 6) = R(t = 6) = e^{-0.32\times 6} = 0.1466.$$

8.2 Reliability of Weibull Distribution Systems

The distribution is named for its developer, Waloodi Weibull, who developed it to describe the observed strengths of tensile test specimen [3]. The Weibull distribution describes the failure times of components when their failure rate either increases or decreases with time. It has the parameters α andβ. Its formula is given by $f(t) = \alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}}, t > 0, \alpha > 0, \beta > 0$ and it follows that the reliability function associated with the Weibull failure-time distribution is given by

$$R\ (t) = e^{-\alpha t^{\beta}} \tag{35-a}$$

The failure rate leading to the Weibull distribution is given by

$$Z\ (t) = \alpha\beta t^{\beta-1}$$

Then $$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\alpha\beta t^{-1} e^{-\alpha t\beta}}{e^{-\alpha t\beta}} \tag{35-b}$$

$$\alpha\beta t^{\beta-1} = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \quad \text{Where } \theta^{\beta} = \frac{1}{\alpha}$$

Thus, we have

$$\text{MTTF} = \text{E (T)} = \theta\Gamma\left(1+\frac{1}{\beta}\right) \tag{36-a}$$

Mean time to failure Weibull model is MTTF.

Variance of the Weibull model is

$$\text{Var (T)} = \sigma_t^{\ 2} = \theta^2\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \left[\Gamma\left(1+\frac{1}{\beta}\right)^2\right]\right\}$$

Now $\qquad \text{R (t/T}_0) = \dfrac{R(t+T_0)}{R(T_0)}$

=

$$\frac{\exp\left[-\left(\frac{t+T_0}{\theta}\right)^{\beta}\right]}{\exp\left[-\left(\frac{T_0}{\theta}\right)^{\beta}\right]} = \exp\left[-\left(\frac{t+T_0}{\theta}\right)^{\beta} + \left(\frac{T_0}{\theta}\right)^{\beta}\right] \tag{36-b}$$

The Weibull distribution is very widely used in reliability modeling. It has the advantages of flexibility in modeling various types of hazard behavior and of algebraic tractability. In addition, as with any two parameter distribution, it can be made to fit many actual situations reasonably well [3].

Solved Problem:

Problem 8.2.1. For a Defence system having a Weibull failure distribution with a shape parameter of 1.4 and a scale parameter of 550 days find: (i) R (100 days) (ii) MTTF (iii) S.D. (iv) The design life for a reliability of 90%.

Solution: The P. D. F. of the Weibull distribution is given by

$$f(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \exp\left\{-\left(\frac{t}{\theta}\right)^{\beta}\right\}, \qquad t \geq 0.$$

Now $\beta = 1.4 \textit{ and } \theta = 550 \textit{ days}$

(i) $$R(t) = \exp\left\{-\left(\frac{t}{\theta}\right)^{\beta}\right\}$$

$\therefore$ $$R(100) = \exp\left\{-\left(\frac{100}{550}\right)^{1.4}\right\} = 0.9122 \; .$$

(ii) $$MTTF = \theta \times \Gamma\left(1 + \frac{1}{\beta}\right)$$

$$= 550 \times \Gamma\left(1 + \frac{1}{1.4}\right) = 550 \times 0.91057$$

$$= 500.8 \; days$$

(iii) $$Var = \theta^2\left\{\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2\right\}$$

$$= 550^2 \times \left\{\Gamma\left(1 + \frac{2}{1.4}\right) - \left(\Gamma\left(1 + \frac{1}{1.4}\right)\right)^2\right\}$$

$$= 550^2 \times \left\{\Gamma(2.43) - (\Gamma(1.71))^2\right\}$$

$$= 550^{2} \times \left[(1.43 \times 0.8860) - (0.91057)^{2}\right]$$

$$S.D. = 550 \times 0.66174 = 363.96 \; days.$$

$$(iv) \; Let \; t_D \; be \; the \; required \; design \; life \; for \; R = 0.90$$

$$\therefore \qquad R(t_D) = \exp\left\{-\left(\frac{t_D}{550}\right)^{1.4}\right\} = 0.90$$

$$\Rightarrow \qquad \left(\frac{t_D}{550}\right)^{1.4} = 0.10536$$

$$\Rightarrow \qquad t_D = 550 \times (0.10536)^{1/1.4} = 110.2 \; days$$

Unsolved Numerical Problems:

Problem 8.2.1. A pressure gauge has a Weibull failure distribution with a shape parameter of 2.1 and a characteristic life of 12,000 hours. Find (i) R(5000) (ii) MTTF (iii) the probability of failure in the first year of continuous operation.

Problem 8.2.2. A Jet engine consists of 5 modules (connected in series) each of which were found to have a Weibull failure distribution with a shape parameter of 1.5. Their characteristics lives are (in operating cycles) 3600, 7200, 5850, 4780 and 9300. Find the reliability function of the engine and the MTTF.

[Ans: $R(t)=e^{-(t/184.7)1.5}$, MTTF=1664.5 cycles]

8.3 Reliability of Normal Distribution Systems

Another popular model of device life length is provided by the Normal distribution. It is a very appropriate model for the reliability evaluation of structural components. If the time to failure T follows a normal distribution N (μ,σ) its P. D. F. is given by,

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(t-\mu)^2}{2\sigma^2}\right], -\infty < t < \infty. \qquad (37)$$

In this case, MTTF = E (T) = μ and

$$\text{Var (T)} = \sigma_T^2 = \sigma^2.$$

R(t) = $\int_t^{\infty} f(t)$ dt is found out by expressing the integral in terms of the standard normal integral and using the normal tables.

Then $\lambda(t) = \dfrac{f(t)}{R(t)}$ is called the instantaneous failure rate or hazard function of the component and the conditional reliability is

$$R(t/T_0) = P\left\{T > \frac{T_0 + t}{T} > T_0\right\} = e^{-\int_{T_0}^{T_0} \lambda(t)dt} \tag{37}$$

Unsolved Problems

Problem 8.3.1. A cutting tool wears out with a time to failure that is normally distributed. It is known that about 34.5% of the tools fail before 9 working days and about 78.8% fail before 12 working days.

(i) Compute MTTF. (Ans. μ=10, σ=2.5, MTTF=10 days)

(ii) Determine its design life for a reliability of 0.99. (Ans. T=4.2 days)

(iii) Determine the probability that the cutting tool will last one more day given that it has been in use for 5 days. (Ans. R = 0.9672)

Problem 8.3.2. A lathe cutting tool has a life time that is normally distributed with an S.D. of 12.0 (cutting) hours. If a reliability of 0.99 is desired over 100 hours of use, find the corresponding MTTF. If the reliability has to life in (0.8, 0.9), find the range within which the tool has to be used? (Ans. 128 years, 112.6 years, 117.9 years)

Problem 8.3.3. An integrated circuit chip has a constant failure rate of 0.02 per thousand hours.

(i) What is the probability that it will operate satisfactorily for at least 20,000 hours? (Ans. 0.6703)

(ii) What is the 5,000 hours reliability of a component consisting of four such ships connected in series? (Ans. 0.6703)

8.4 Reliability and MTBF Evaluation Using Exponential Model

The most widely used distribution function for modeling reliability is the exponential distribution. Reliability and MTBF of series and parallel connected Defence support systems can be evaluated using Exponential model as follows.

8.4.1 Series Connected Components

Suppose now that a system consists of n components connected in series and that these components have the respective failure rates $\alpha_1, \alpha_2, \ldots \alpha_n$. The product law of reliabilities can be written as

$$R_s(t) = e^{-t\sum_{i=1}^{n}\alpha_i} \tag{38}$$

The mean time between failures (MTBF) of a series system is,

$$\mu_s = \frac{1}{\frac{1}{\mu_1} + \frac{1}{\mu_2} + \ldots\ldots + \frac{1}{\mu_n}} \tag{39}$$

Where, μ_i is MTBF of *i*th component. In the special case where all n components have the same failure rate α and hence the same MTBF μ, the system failure rate is $n\alpha$, and the system MTBF is $\frac{1}{\mu\alpha} = \frac{\mu}{n}$.

8.4.2 Parallel Connected Components

The mean time to failure of a parallel system is also difficult to obtain in general, but in the special case where all components have the same failure rate α, an interesting and useful result can be obtained. In this special case the system reliability function becomes,

$$R_p(t) = 1 - (1 - e^{-\alpha t})^n = \binom{n}{1} e^{-\alpha t} - \binom{n}{2} e^{-2\alpha t} + \ldots\ldots + (-1)^{n-1} e^{-n\alpha t}$$

Then, making use of the fact that, $f_p(t) = R'_p(t)$, We obtain, $f_p(t) = \alpha\begin{pmatrix} n \\ 1 \end{pmatrix} e^{-\alpha t} - 2\alpha\begin{pmatrix} n \\ 2 \end{pmatrix} e^{-2\alpha t} + \ldots\ldots + (-1)^{n-1} n\alpha e^{-n\alpha t}$ (40)

and the mean of the failure-time distribution is given by, $\mu_p = \int_0^\infty t.f_P(t)dt$

$$= \alpha\begin{pmatrix} n \\ 1 \end{pmatrix}\int_0^\infty te^{-\alpha t}dt - 2\alpha\begin{pmatrix} n \\ 2 \end{pmatrix}\int_0^\infty t\,e^{-2\alpha t}\,dt + \ldots\ldots + (-1)^{n-1} n\alpha\int_0^\infty te^{-n\alpha t}dt$$

$$= \frac{1}{\alpha}\binom{n}{1} - \frac{1}{2\alpha}\binom{n}{2} + \ \ldots. \ + \ (-1)^{n-1}\frac{1}{n\alpha}$$

Then mean time between failures (MTBF) in parallel system is $\mu_p = \frac{1}{\alpha}\left(1 + \frac{1}{2} + \ldots\ldots\ldots\ldots + \frac{1}{n}\right)$. The MTBF of the system consists of n components having the identical failure rate α provided each defective component is replaced whenever the whole parallel system fails. Thus, if we use two parallel components rather than one, the mean time to failure of the pair exceeds that of the single component by 50 percent, rather than doubling it.

Unsolved Problem:

Problem 8.4.1. An airlines maintains an online reservation system with a standby computer available if the primary fails. The on-line system fails at the constant rate of once per day while the standby fails (only when online) at the constant rate of twice per day. If the primary unit may be repaired at a constant rate with an MTTR of 0.5 of a day, what is the single day reliability? (Ans. 0.7125)

8.5 Miscellaneous Solved Numerical Problems

Problem 8.5.1 In a Defence support system, given that R (t) $= 1/e^{\sqrt{0.001t}}, t \geq 0$

(i) Compute the reliability for a 50 hours mission

(ii) Show that the hazard rate is decreasing

(iii) Given a 10 hour wear-in period, compute the reliability for 50 hour mission.

(iv) What is the design life for a reliability of 0.95, given a 10 hour wear-in period?

Solution: $R(t) = 1/e^{\sqrt{0.001t}}\ ,t \geq 0$

(i) $R(50) = 1/e^{\sqrt{0.001\times 50}} \qquad = 0.9512$

(ii) $\lambda(t) = \dfrac{-R'(t)}{R(t)} = -e^{\sqrt{0.001t}} \times \dfrac{1}{e^{\sqrt{0.001t}}} \times \left(-\sqrt{\dfrac{0.001}{t}}\right)$

$$= \sqrt{\frac{0.001}{t}}, which\ is\ a\ decrea\sin g\ \ function\ \ of\ \ t.$$

(iii) $R(t/T_0) = \dfrac{R(T_0 + t)}{R(T_0)}$

$\therefore$ $R(50/10) = \dfrac{R(60)}{R(10)} = \dfrac{1}{e^{\sqrt{0.001\times 60}}} \times e^{\sqrt{0.001\times 10}} = 0.8651$

(iv) $R(t_D + 10) = 0.95$

i.e. $\dfrac{R(t_D + 10)}{R(10)} = 0.95 \qquad \Rightarrow \qquad R(t_D + 10) = R(10) \times 0.95$

$\Rightarrow$ $1/e^{\sqrt{0.001\times (t_D + 10)}} = 0.95 \times 1/e^{\sqrt{0.001\times 10}}$

$\Rightarrow$ $\sqrt{0.001\times (t_D + 10)} = 0.15129 \ \ \Rightarrow \qquad t_D = 12.89\, hours$

Problem 8.5.2. A relay circuit has an MTBF of 0.8 year. Assuming random failures

(i) Calculate the probability that the circuit will survive 1 year without failure.

(ii) What is the probability that there will be more than 2 failures in the first year?

(iii) What is the expected number of failure per year?

Solution: Since the failures are random events, the number of failures in an interval of length t follows a Poisson process, given by P[N(t) = n]= $\frac{e^{-\lambda t}(\lambda t)n}{n!}, n \geq 0$, where λ = failure rate.

Then the time between failures follows an exponential distribution with mean 1/λ.

Now $$MTBF = \frac{1}{\lambda} = 0.8 \ year$$

$\therefore$ $$\lambda = \frac{1}{0.8} per\, year = 1.25 \ per \ year$$

(i) $$P[N(1) = 0] = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-1.25} = 0.2865$$

(ii) $$P[N(1) > 2] = 1 - e^{-\lambda}[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!}] = 1 - 0.2865\left[1 + 1.25 + \frac{(1.25)^2}{2}\right] = 0.1315.$$

(iii) $$E[N(t)] = \lambda t$$

$$\therefore \quad E\{number of \ failure\, per\, year\} = \lambda = 1.25.$$

Problem 8.5.3. For a redundant Defence support system with n independent identical components with constant failure rate λ Show that MTTF is equal to $\frac{1}{\lambda}\sum_{i=1}^{n} {}^nC_1 \frac{(-1)^{i-1}}{i}$, If λ = 0.02 per hour, what is the minimum value of the system reliability,

Solution: If $R_s(t)$ is the system reliability,

$R_s(t) = 1 - (1 - e^{-\lambda t})^n$, since the component reliability = $e^{-\lambda t}$

$$= 1 - \sum_{i=0}^{n} (-1)^i \ {}^nC_i e^{-i\lambda t} = \sum_{i=1}^{n} (-1)^{i-1} \ {}^nC_i e^{-i\lambda t}$$

$$\therefore \qquad MTTF = \int_0^\infty R_s(t)dt$$

$$= \sum_{i=1}^{n} (-1)^{i-1} \; {}^nC_1 \int_0^\infty e^{-i\lambda t}\, dt = \frac{1}{\lambda}\sum_{i=1}^{n} {}^nC_i \frac{(-1)^{i-1}}{i}$$

$$When \;\; n = 2, \qquad MTTF = 200 \times \sum_{i=1}^{2} {}^2C_i \frac{(-1)^{i-1}}{i} = 200\left[2 - \frac{1}{2}\right] = 300$$

$$When \;\; n = 3, \qquad MTTF = 200 \times \sum_{i=1}^{3} {}^3C_i \frac{(-1)^{i-1}}{i} = 200\left(3 - \frac{3}{2} + \frac{1}{3}\right) = 200 \times \frac{11}{6} = 366$$

$$Hence\ the\ required\ \min imum\ value\ of\ n = 4.$$

Problem 8.5.4. Six identical components with constant failure rates are connected in (i) high level redundancy with 3 components in each subsystem (ii) low level redundancy with 2 components in each subsystem. Determine the component MTTF in each case, necessary to provide a systems reliability of 0.90 after 100 hours of operation.

Solution: Let λ be the constant failure rate of each component.

Then $R = e^{-\lambda t}$, for each component.

(i) For high level redundancy,

$$R_s(t) = 1 - [1 - \{R(t)\}^3]^2 = [1 - (1 - e^{-3\lambda t})^2]$$

$$\therefore \qquad R_s(100) = [1 - (1 - e^{-300\lambda})^2] = 0.90$$

$$\Rightarrow \qquad (1 - e^{-300\lambda})^2 = 0.10 \qquad \Rightarrow \qquad 1 - e^{-300\lambda} = 0.3162$$

$$\Rightarrow \qquad e^{-300\lambda} = 0.6837$$

$$\therefore \qquad 300\lambda = 0.3801 \qquad \Rightarrow \qquad \lambda = \frac{0.3801}{300}$$

$$\therefore \text{ MTTF of each component } \frac{1}{\lambda} = \frac{300}{0.3801} = 789.2 \;\; hours.$$

(ii) For low level redundancy,

$$R_s(t) = [1 - \{1 - [R(t)]\}^2]^3 = [1 - \{1 - e^{-\lambda t}\}^2]^3$$

$$\therefore \quad R_s(100) = [1 - \{1 - e^{-100\lambda}\}^2]^3 = 0.90$$

$$\Rightarrow \quad 1 - (1 - e^{-100\lambda})^2 = 0.9654 \quad \Rightarrow \quad (1 - e^{-100\lambda})^2 = 0.345$$

$$\Rightarrow \quad 1 - e^{-100\lambda} = 0.1857 \quad \Rightarrow \quad e^{-100\lambda} = 0.8142$$

$$\Rightarrow \quad 100\lambda = 0.20551$$

$\therefore$ MTTF of each component = $\dfrac{1}{\lambda} = \dfrac{100}{0.20551} = 486.6$ hours.

Problem 8.5.5. A computer has an MTTF = 34 hours and an MTTR = 2.5 hours. What is the steady-state availability? If the MTTR is reduced to 1.5 hours, what MTTF can be tolerated without decreasing the steady-state availability of the computer?

Solution: Steady-state availability = $\dfrac{MTTF}{MTTF + MTTR} = \dfrac{34}{34 + 2.5} = 0.931$

MTTR is reduced to 1.5 hour.

Now, $$\frac{MTTF}{MTTF + 1.5} = 0.931$$

$$\Rightarrow \quad \text{MTTF} = 0.931\ (\text{MTTF} + 1.5)$$

$$\Rightarrow \quad (1 - 0.931)\ \text{MTTF} = 0.931 \times 1.5$$

$$\Rightarrow \quad MTTF = \frac{0.931 \times 1.5}{1 - 0.931} = 20.239 \ \ hours.$$

Problem 8.5.6. An engine health monitoring Defence system consists of a primary unit and a stand by unit. The MTTF of the primary unit is 1000 operating hours and the MTTF of the stand by unit is 333 hours when in operation. There are no failures while the backup unit is in stand by. If the primary unit may be repaired at a repair rate of 0.01 per hour, while the stand by unit is operating, estimate the design life for a reliability of 0.90.

Solution: For the stand by redundant system,

$$R(t) = \frac{m_1}{m_1 - m_2} e^{m_2 t} - \frac{m_2}{m_1 - m_2} e^{m_1 t} \qquad \ldots (1)$$

Where m1, m2 are roots of the equation

$$m^2 + (\lambda_1 + \lambda_2 + \mu)\, m + \lambda_1 \lambda_2 = 0 \qquad \ldots (2)$$

Here

$$\lambda_1 = \frac{1}{1000} = 0.001 / hour \quad ; \quad \lambda_2 = \frac{1}{333} = 0.003 / hour \quad and \quad \mu = 0.01 / hour$$

Using these values in (2), we have

$$m^2 + 0.014\, m + 0.000003 = 0$$

$$\therefore m_1, m_2 = \frac{-0.014 \pm \sqrt{(0.0-14)^2 - 4 \times 0.000003}}{2} = (0.00022, -0.01378)$$

Using these values in (1), we get

$$R(t) = -0.01622 \times e^{-0.01378t} + 1.01622 \times e^{-0.00022\, t}$$

When the reliability is 0.90, the design life D is given by

$$1.01622 \times e^{-0.00022\, t\, D} - 0.01622 \times e^{-0.01378\, D} = 0.90$$

Solving this equation by trials, we get.

D = 550 hours

Problem 8.5.7. A critical communications relay has a constant failure rate of 0.2 per day once it has failed, the mean time to repair is 2.5 days (the repair rate is constant).

(i) What are the point availability at the end of 2 days, the interval availability over a 2 day mission, starting from zero and the steady-state availability?

(ii) If two communication relays operate in series, compute the availability at the end of 2 days.

(iii) If they operate in parallel, compute the steady-state availability of the system.

(iv) If one communication relay operates in a stand by made with no failure in stand by, what is the steady-state availability?

Solution: Here $\lambda = 0.2\, per\, day; \frac{1}{\mu} = 2.5 \therefore \mu = 0.4\, per\, day$

(i) The point availability

$$A_P(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)}\, t$$

$$\therefore \qquad A_P(2) = \frac{0.4}{0.2+0.4} + \frac{0.2}{0.2+0.4} e^{-(0.2+0.4)\times 2} = 0.66 + (0.33 \times 0.30) = 0.7590$$

$$A_I(T) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{(\lambda+\mu)^2 \times T}\{1 - e^{-(\lambda+\mu)T}\}$$

$$\therefore \qquad A_I(T) = \frac{0.4}{0.2+0.4} + \frac{0.2}{(0.2+0.4)^2 \times 2}\{1 - e^{-(0.2+0.4)\times 2}\}$$

$$0.66 + 0.28 \times (1 - 0.30) = 0.8560$$

$$A\ (\infty) = \frac{\mu}{\lambda+\mu} + \frac{0.4}{0.2+0.4} = 0.66$$

(ii) $\qquad A_S\ (2) = \{A_p(2)\}^2 = (0.7590)^2 = 0.5760.$

(iii) $\qquad A_S(\infty) = 1 - \{1 - A(\infty)\}^2 = 1 - \{1 - 0.66\}^2 = 0.84$

(iv)For the standby redundant system

$$A_S(\infty) = \frac{\lambda_1 \mu + \mu^2}{\mu^2 + \lambda_1 \mu + \lambda_1 \lambda_2}; \quad Here\ \lambda_1 = \lambda_2 = 0.2\ \ and\ \ \lambda = 0.4$$

$$A_S(\infty) = \frac{(0.2)(0.4) + (0.4)^2}{(0.4)^2 + (0.2)(0.4) + (0.2)(0.2)} = 0.8571.$$

Problem 8.5.8. A pipe line processor possesses five serial segments, each of which must process data interns to produce the final output of the processor. If each segment possesses a constant failure rate of 0.23 failures per year, segment failures being independent, find the reliability function and MTTF for the system.

Solution: Here failure rate is constant.

Then the failure time distribution is exponential with parameter $\lambda = 0.23.$

$$\therefore \qquad R_1(t) = e^{-(0.23)t} for \ \ Onesegment$$

Here five segments are in series.

Since the segment failures being independent, reliability for the system is

$R(t) = R_1(t).\ R_2(t).\ R_3(t).\ R_4(t).\ R_5(t)$

$$= e^{-(0.23)t} \times e^{(0.23)t} \times e^{-(0.23)^t} \times e^{-(0.23)^t} \times e^{-(0.23)^t} = e^{-(1.15)\ t}$$

$$\text{MTTF} = \int_0^\infty R(t)dt = \int_0^\infty e^{-(1.15)} dt = \left[\frac{e^{-(1.15)t}}{-1.15}\right] = 0.8696$$

Problem 8.5.9. A new machine has a constant failure rate of 0.03 per day (assuming continuous use) and a constant repair rate of 0.2 per day.

(a) Compute the interval availability for the first 30 days and the steady state availability.

(b) Determine the steady-state availability if a stand by unit is purchased. Assume no failures in standby.

(c) If both units are active, what is the steady-state availability?

Solution: λ=0.03 per day, μ=0.2 per day

(a)

$$A_1(T) + \frac{\mu}{\lambda+\mu} + \frac{\mu}{(\lambda+\mu)^2 \times T}[1 - e^{-(\lambda+\mu)T}]$$

$$\therefore \qquad A_1(30)\frac{0.2}{0.03+0.2} + \frac{0.03}{(0.03+0.2)^2 \times 30}[1 - e^{-(0.03x0.2)x30}]$$

$$= 0.8696 + 0.0031 + 0.8727$$

$$A(\infty) = \frac{\mu}{\lambda+\mu} = 0.8696$$

(b) For the stand by redundant system

$$A_s(\infty) = \frac{\lambda\mu + \mu^2}{\mu^2 + \lambda\mu + \lambda^2} = \frac{(0.03\times 0.2) + (0.2\times 0.2)}{(0.2)^2 + (0.03\times 0.2) + (0.03)^2}$$

$$=> \frac{0.046}{0.0469} = 0.9808\ .$$

(C) For the active redundant system

$A_i(\infty)=1-[1-A(\infty)]^2$ => $1-[1-0.8696]^2$ =0.9830 Ans.

9 Conclusion

The analysis presented in this chapter provides a means for relating system reliability to component reliability for many types of equipment designs. The system configurations based on binary component states and independent components are sufficient to permit a reductionist approach to reliability analysis. Reliability should be studied at the component level because the dependence of system reliability on component reliability is well defined. For Defence support system designs, the ability to focus independently on individual component reliability performance is essential to achieving the high levels of reliability. Examples on applications of the methods are given to illustrate the advantages and limitations of the different techniques, together with case studies drawn from the author's experience of academia and consultancy. Comprehensive coverage of the basic concepts of probability theory, IDSS structures with reliability evaluations, Hazard model for failure analysis and various probability distributions of IDSS, solved as well as unsolved numerical examples based on IDSS in each sub-section have been described in this chapter.

Acknowledgements

Author is grateful to Department of Electrical Engineering, Institute of Technology, Banaras Hindu University; Varanasi (Uttar Pradesh); India for providing encouragements and necessary facilities to complete this chapter in time. He dedicates this valuable work to sweet memories of his great father Parinibutta Ram Das Saket Vardhan who has passed away on August 17, 2008. Without his heartily blessings, constant supports and timely encouragements it would not have been possible to complete this scientific, academic and technical work for Springer-Verlag publishers, Germany in its present shape. He would also like thank to Aanchal, Aakanksha, Siddhartha, and Siddhanta Vardhan for providing biorhythmic emotional and intellectual supports during the writing work of the chapter.

References

[1] Billinton, R., Allan, R.N.: Reliability Evaluation of Power Systems, 2nd edn. Plenum, New York (1996)

[2] Hunter, R.E., Donna, F.: European Security and Defense Policy: NATO's Companion or Competitor. RAND Corporation publishers (2002) ISBN-13: 9780833031174, ISBN: 083303117

[3] Nachlas, J.A.: Reliability Engineering: Probabilistic Models and Maintenance Methods. CRC Pr I Llc publishers (July 2005); 10 digit ISBN: 0849335981 and 13 digit ISBN: 9780849335983

[4] Saket, R.K.: Reliability Evaluation of Interconnected Composite Electrical Power Systems Considering Voltage Stability and Continuation Power Flow. Ph. D. Thesis, Faculty of Engineering and Technology, Sam Higginbottom Institute of Agriculture, Technology and Sciences, Allahabad (Uttar Pradesh), India (2006)

[5] Saket, R.K.: Design, Development and Reliability Evaluation of Micro Hydro Power Generation System Based on Municipal Waste Water. In: Proceedings of the IEEE Electrical Power and Energy Conference (EPEC 2008), IEEE Canada and IEEE Vancouver Section, The University of British Columbia, Vancouver, BC Canada, V6T1Z4, British Columbia, October 6-7, pp. 1–8 (2008), `http://www.ieeexplore.ieee.org`

[6] Ireson, W.G., Coombs, C.F., Moss, R.Y.: Handbook of Reliability Engineering and Management, 2nd edn. (Published 1995)

[7] O'Connor, P.D.T.: Practical Reliability Engineering, 4th edn. (Published 2002), `http://www.amazon.com/exec/obidos`

[8] Saket, R.K., Kaushik, (Wg. Cdr.) S.P., Singh, G.: Probabilistic Approach for Biorhythmic Analysis to prevent Aviation Accidents. International Journal of Intelligent Defence Supports Systems 1(4), 272–289 (2008), `http://www.inderscience.com/ijidss`

[9] Saket, R.K., Kaushik, (Wg. Cdr.) S.P.: Use of Gaussian Distribution Method for Biorhythmic Analysis to Prevent Aviation Accident. In: Proceedings of the 14th ISME International Conference on Mechanical Engineering in Knowledge age, New Delhi, India, pp. 139–147 (December 2005)

[10] Kaushik, S.P., Murthy, W.S., Shrivastava, K.K.: Biorhythmic approach for accident prevention in military aviation. Defence Institute of Physiology and Allied Sciences, New Delhi, India, Project Report AR & DB, No: 820 (1990)

[11] Thompson, S.G., Ghanea-Hercock, R. (eds.): DAMAS 2005. LNCS (LNAI), vol. 3890. Springer, Heidelberg (2006) ISBN: 978-3-540-32832-2

[12] Grant Ireson, W., Coombs, C.F., Moss, R.Y.: Handbook of Reliability Engineering and Management. McGraw-hill Professional Publishing, New York, ISBN: 0070127506, ISBN-13:9780070127500, 978-0070127500

[13] Clarke, M. (ed.): Brassey's Defence Yearbook, 1995. Centre for Defence Studies, Publisher, Brassey's (U.K.) Ltd (1995), `http://www.alibris.co.uk/search/books/` ISBN-13: 9781857531312 ISBN: 1857531310

[14] Reeve, D.: Risk and Reliability: Coastal and Hydraulic Engineering. Spon Press (August 13, 2009) ISBN: 978-0-415-46755-1

[15] Arya, L.D., Chaube, S.C., Saket, R.K.: Generation system adequacy evaluation using probability theory. Journal of the Institution of Engineers (India) 81, 170–174 (2001), `http://www.ieindia.org`

[16] Saket Dr., R.K., Bansal Dr., R.C., Singh, G.: Generation Capacity Adequacy Evaluation based on Peak Load Considerations. The South Pacific International Journal on Natural Sciences, University of the South Pacific, Suva, Fiji Island 24, 38–44 (2006), `http://www.fst.usp.ac.fj`, `http://www.fst.usp.ac.fj`

[17] Leemis, L.M.: Reliability: Probabilistic Models and Statistical Methods (Published 1995), `http://www.amazon.com/exec/obidos`

[18] Wesley Barnet, J.: Statistical Analysis for Engineers and Scientists: A Computer-Based Approach (Published 1994), `http://www.amazon.com/exec/obidos`

[19] Meeker, W.K., Escobar, L.A.: Statistical Methods for Reliability Data. Wiley Series in Probability and Statistics; Applied Probability and Statistics Section (Published 1998), `http://www.amazon.com/exec/obidos`

[20] Rigdon, S.E., Basu, A.P.: Statistical Methods for the Reliability of Repairable Systems (Published 2000), `http://www.amazon.com/exec/obidos`

[21] Saket, R.K., Bansal, R.C., Singh, G.: Power System Compo-nents Modelling and Reliability Evaluation of Generation Capacity. International Journal of Reliability and Safety 3(4), 427–441 (2009), `http://www.inderscience.com/ijrs`

[22] Kececioglu, D.: Reliability and Life Testing Handbook, vol. 1 and 2 (Published 1994), `http://www.reliasoft.com/books/kcbooks/`

[23] Singh, K.B., Saket, R.K., Chatterji, S.: The effect of the derating on component selection and product design. International Journal of Reliability and Safety 4(1), 89–103 (2010), `http://www.inderscience.com/ijrs`

[24] Nelson, W.: Accelerated Testing: Statistical Models, Test Plans and Data Analyses. Wiley Series in Probability and Mathematical Statistics-Applied Probability (Published 1990), `http://www.amazon.com/exec/obidos`

[25] Kececioglu, D., Sun, F.-B.: Burn-In Testing: Its Quantification and Optimization (Reliability Engineering) (Published 1997), `http://www.reliasoft.com/books/kcbooks/`

[26] Nelson, W.: Recurrent-Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications. (Published 2003), `http://www.amazon.com/exec/obidos`

[27] Saket, R.K., Bansal, R.C.: Generating Capacity Reliability Evaluation Using Gaussian Distribution Approach. In: Proceedings of the International Conference on Reliability and Safety Engineering, Reliability Engineering Centre, Indian Institute of Technology, Kharagpur (West Bengal), India, December 21-23, pp. 261–274 (2005), `http://www.iitkgp.ac.in`

[28] Saket, R.K., Zobaa, A.F.: Power System Reliability Evaluation Using Safety Factor Concept and Simpsons 1/3rd rule. In: International Federation of Automatic Control (IFAC) Symposium on Power Plants and Power System Control (PPPSC), Tampare Hall, Finland (United Kingdom), July 5-8 (2009), `http://wwwautomaatioseura.fi/confprog/conference`

[29] Kececioglu, D.: Maintainability, Availability and Operational Readiness Engineering Handbook (Published 1995), `http://www.reliasoft.com/books/kcbooks/`

[30] Kececioglu, D.: Reliability and Life Testing Handbook, vol. 1 (Published 1993), `http://www.reliasoft.com/books/kcbooks/`

Author Index

Aidman, Eugene 1

Butavicius, M.A. 183
Butavicius, Marcus 147

Dasgupta, Prithviraj 5

Foster, R. 183

Graves, Ian 147

Heyer, Rebecca 147

Ivancevic, Vladimir G. 21

Jacques, Philip 147
Jain, Lakhmi C. 1
Johnson, Raymond 147

Kaushik, Wg. Cdr. S.P. 207
Kuester, Natalie 147

MacLeod, V. 183
MacLeod, Veneta 115, 147
McCormac, A. 183
McCormac, Agata 147
McLindin, Brett 115

Parsons, K.M. 183
Parsons, Kathryn 147

Reid, Darryn J. 21

Saket, R.K. 207, 241
Singh, Col. Gurmit 207

Whittenbury, A. 183