Syntactic Measures of Complexity

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Arts

1999

Bruce Edmonds

Department of Philosophy

Table of Contents

Table of Contents - page 2

List of Figures - page 11

List of Tables - page 13

Abstract - page 14

Declaration - page 15

Notes of copyright and the ownership of intellectual property rights - page 15

The Author - page 16

Acknowledgements - page 16

1 - Introduction - page 17

1.1 - Background - page 17

1.2 - The Style of Approach - page 18

1.3 - Motivation - page 19

1.4 - Style of Presentation - page 20

1.5 - Outline of the Thesis - page 21

2 - Models and Modelling - page 23

2.1 - Some Types of Models - page 25

2.2 - Combinations of Models - page 28

2.3 - Parts of the Modelling Apparatus - page 33

2.4 - Models in Machine Learning - page 38

2.5 - The Philosophical Background to the Rest of this Thesis - page 41

3 - Problems and Properties - page 44

3.1 - Examples of Common Usage - page 44

3.1.1 - A case of nails - page 44

3.1.2 - Writing a thesis - page 44

3.1.3 - Mathematics - page 44

3.1.4 - A gas - page 44

3.1.5 - An ant hill - page 45

3.1.6 - A car engine - page 45

3.1.7 - A cell as part of an organism - page 46

3.1.8 - Computer programming - page 46

3.2 - Complexity as a Comparison - page 46

3.2.1 - The emergence of life - page 47

3.3 - What the Property of Complexity Could Usefully Refer to - page 47

3.3.1 - Natural systems - page 47

3.3.2 - The interaction of an observer with a system - page 53

3.3.3 - Patterns - page 55

3.3.4 - The modelling relation - page 56

3.3.5 - A model with respect to a specified framework - page 56

3.4 - Some Unsatisfactory Accounts of Complexity - page 57

3.4.1 - Size - page 57

3.4.2 - Size of rules - page 58

3.4.3 - Minimal size - page 58

3.4.4 - Processing time - page 59

3.4.5 - Ignorance - page 60

3.4.6 - Variety - page 61

3.4.7 - Midpoint between order and disorder - page 62

3.4.8 - Improbability - page 63

3.4.9 - Expressivity - page 65

3.4.10 - Dimension - page 65

3.4.11 - Ability to surprise - page 66

3.4.12 - Logical strength - page 66

3.4.13 - Irreducibility - page 67

3.5 - Complexity is Relative to the Frame of Reference - page 68

3.5.1 - The level of application - page 68

3.5.2 - Goals - type of difficulty - page 69

3.5.3 - Atomic parts - page 69

3.5.4 - The language of description - page 69

4 - A Definition of Complexity - page 72

4.1 - Aspects of the Definition - page 75

4.1.1 - Identity of a system - page 75

4.1.2 - Atomic components - page 76

4.1.3 - Difficulty - page 77

4.1.4 - Formulating overall behaviour - page 78

4.1.5 - Complexity vs. ignorance - page 79

4.1.6 - As a gap between the global and local - page 80

4.1.7 - The comparative nature of complexity - page 80

4.1.8 - The existence of complexity - page 81

4.1.9 - Relativisation to a language - page 81

4.2 - Examples - page 81

4.2.1 - The flight behaviour of a herd - page 82

4.2.2 - Cellular automata - page 82

4.3 - Relationship to Some Other Formulations - page 83

4.3.1 - Number of inequivalent descriptions - page 83

4.3.2 - Effective measure complexity - page 84

4.3.3 - Computational complexity - page 84

4.3.4 - Algorithmic information complexity - page 84

4.3.5 - Shannon entropy - page 85

4.3.6 - Crutchfield’s “topological complexity” - page 85

5 - Applications of Complexity to Formal Languages - page 86

5.1 - Types of Complexity Involving Formal Languages - page 86

5.2 - Expected Properties of “Analytic Complexity” - page 87

5.2.1 - Independent of the particular symbols used - page 87
5.2.2 - The complexity of sub-expressions should be less than the whole - page 87
5.2.3 - Expressions with no repetitions are simple - page 88
5.2.4 - Small size should limit the possible complexity - page 89
5.2.5 - There should be no upper limit to complexity if the language is suitably generative - page 89
5.2.6 - The complexity of irrelevant substitutions - page 90
5.2.7 - The complexity of relevant relating of expressions - page 91
5.2.8 - Decomposability of expressions - page 91
5.3 - Measures of Analytic Complexity - page 93
5.3.1 - Notation - page 93
5.3.2 - Weak complexity measures - page 95
5.3.3 - Weak complexity measures where simple repetition does not increase complexity - page 100
5.3.4 - Weak complexity measures that respect the subformula relation and where simple repetition does not increase complexity - page 102
5.3.5 - Strong complexity measures - page 104
5.4 - The Cyclomatic Number as a Measure of Analytic Complexity - page 106
5.5 - Layers of Syntax and Complexity - page 108
5.5.1 - Example 1 - a supply of variable names - page 109
5.5.2 - Example 2 - WFFs of the implication fragment - page 110
5.5.3 - Example 3 - The implicational fragment of E - page 111
5.5.4 - Discussion of syntactic structures - page 112
5.6 - Application to Axioms and Proof Systems - page 113
5.6.1 - Axiom complexity - page 113
5.6.2 - Proof complexity - page 117
5.7 - Application to Simplification - page 120
5.7.1 - Searching over equivalent expressions within a language - page 120
5.7.2 - Searching over equivalent derivations within a language - page 121
5.7.3 - Specialising the syntactic level - page 121

5.7.4 - Searching over equivalent languages - page 123

5.7.5 - Simplification via trade-offs with specificity and accuracy - page 124

6 - Philosophical Applications - page 126

6.1 - Complexity and Relevance - page 126

6.2 - Complexity and Emergence - page 126

6.3 - Complexity and Language - page 128

6.4 - Complexity and Representation - page 128

6.5 - Complexity and “Simplicity” - page 129

6.6 - Complexity and Evolution - page 130

6.7 - Complexity and Holism - page 131

6.8 - Complexity and System Identity - page 132

6.9 - Complexity and Society - page 133

7 - Conclusion - page 134

7.1 - Further Work - page 134

8 - Appendix 1 - A Brief Overview of Some Existing Formulations of Complexity - page 136

8.1 - Abstract Computational Complexity - page 136

8.2 - Algorithmic Information Complexity - page 136

8.3 - Arithmetic Complexity - page 138

8.4 - Bennett's ‘Logical Depth’ - page 138

8.5 - Cognitive Complexity - page 139

8.6 - Connectivity - page 140

8.7 - Cyclomatic Number - page 140

8.8 - Descriptive/Interpretative Complexity - page 141

8.9 - Dimension of Attractor - page 141

8.10 - Ease of Decomposition - page 142

8.11 - Economic Complexity - page 142

8.12 - Entropy - page 143

8.13 - Goodman's Complexity - page 143

8.14 - Horn Complexity - page 143

8.15 - Information - page 144

8.16 - Information Gain in Hierarchically Approximation and Scaling - page 145

8.17 - Irreducibility - page 145

8.18 - Kemeny's Complexity - page 146

8.19 - Length of Proof - page 146

8.20 - Logical Complexity/Arithmetic Hierarchy - page 146

8.21 - Loop Complexity - page 147

8.22 - Low Probability - page 148

8.23 - Minimum Number of Sub Groups - page 148

8.24 - Minimum Size - page 149

8.25 - Mutual Information - page 151

8.26 - Network Complexity - page 151

8.27 - Number of Axioms - page 152

8.28 - Number of Dimensions - page 152

8.29 - Number of Inequivalent Descriptions - page 152

8.30 - Number of Internal Relations - page 153

8.31 - Number of Spanning Trees - page 153

8.32 - Number of States in a Finite Automata - page 153

8.33 - Number of Symbols - page 154

8.34 - Number of Variables - page 155

8.35 - Organised/Disorganised Complexity - page 155

8.36 - Shannon Information - page 156

8.37 - Simplicity - page 156

8.38 - Size - page 157

8.39 - Size of Grammar - page 157

8.40 - Size of matrix - page 158

8.41 - Sober's Minimum Extra Information - page 158

8.42 - Sophistication - page 159

8.43 - Stochastic Complexity - page 160

8.44 - Syntactic Depth - page 160

8.45 - Tabular Complexity - page 161

8.46 - Thermodynamic Depth - page 161

8.47 - Time and Space Computational Complexity - page 162

8.48 - Variety - page 163

9 - Appendix 2 - Longer Proofs - page 164

9.1 - (Non-existence of) Complexity Measures on Strings - page 164

9.2 - Cyclomatic Number as a Lower Bound for Minimal Damage Cut - page 169

9.3 - Decomposition of Formulas into Complexes - page 169

9.4 - Generating a Measure from a Function on the Complexes - page 170

9.5 - Three Conditions that are Equivalent on a Weak Complexity Measure - page 175

10 - Appendix 3 - Formalisation of Syntactic Structure - page 182

10.1 - Formalisation - page 182

10.1.1 - The syntax of trees - page 182

10.1.2 - The syntax of rules - page 182

10.1.3 - The syntax of syntactic structures: - page 182

10.1.4 - Generation from syntactic structures - page 183

10.1.5 - Production from trees - page 183

10.1.6 - Complete production - page 184

10.1.7 - Complete productive generation from syntactic structures - page 184

10.2 - The Expressivity of Syntactic Structures - page 184

10.3 - Flattening Syntactic Structures - page 187

11 - Appendix 4 - A tool for exploring syntactic structures, complexity and simplification - page 188

11.1 - Overview - page 188

11.2 - Examples - page 188

11.3 - Programming - page 190

11.4 - Interface - page 192

12 - Appendix 5 - A comparison of different rankings of logical formula - page 193

13 - Appendix 6 - Complexity and Scientific Modelling - page 199

Overview - page 199

Complexity - page 200

A Framework for Analysing Modelling - page 201

Other Formulations of Complexity - page 202

Order and Disorder - page 203

Noise - page 206

Complexity vs. Information - page 207

Complexity and Induction - page 207

Conclusion - page 208

14 - Appendix 7 - Complexity and Economics - page 210

What is Complexity? - page 210

“Complexity” in economics - page 210

The “Sciences of Complexity” - page 210

Complexity *per se - page 211*

The effects of complexity on modelling by agents - page 213

Ideal rationality and perfect information - page 213

Ideal rationality and noisy information - page 213

Ideal rationality and inadequate information - page 214

Bounded rationality and inadequate information - page 214

The effects of modelling by agents on complexity - page 215

Ignoring the process of modelling by economic agents - page 215
Including the process of modelling by economic agents - page 215
Towards dealing with the complexity of modelling agents - modelling modelling - page 217
The Form - meaning distinction - page 217
The complexity, specificity, error trade-off - page 218
The modelling language - page 219
Processes of model development - page 219
Some future directions for economic modelling - page 220
Applying our model of modelling to ourselves - page 220
Relatively new (non-numerical) techniques - page 221
Conclusion - complexity again - page 222

15 - References - page 223

List of Figures

Figure 1: Entity, W, using A as a model of B - page 24
Figure 2: An illustration of the syntactic view of models - page 28
Figure 3: A semantic picture of modelling (from Giere in [176]) - page 29
Figure 4: The picture with explanatory and phenomenological models - page 30
Figure 5: The picture with the data model added - page 31
Figure 6: The models used in chemical simulation (adapted from [201]) - page 32
Figure 7: Model specification and semantics - page 34
Figure 8: Relating model-structures via the ‘subsumption’ of model contents - page 36
Figure 9: Relating model-structures via approximation of model contents - page 36
Figure 10: Relating model-structures via ‘chaining’ - page 37
Figure 11: The ‘modelling relation’ from [385] - page 41
Figure 12: Observer–system pair - page 53
Figure 13: An observer-system pair with multiple models - page 54
Figure 14: Complete order, chaos and complete disorder - page 62
Figure 15: Presumed graph of disorder against complexity - page 62
Figure 16: Possible diagrammatic inclusions - page 63
Figure 17: Finding a description to cover a data model - page 73
Figure 18: Measuring the difficulty of finding a model description - page 74
Figure 19: A formula represented as a tree - page 107
Figure 20: A formula represented as a collapsed tree - page 108
Figure 21: A single layer syntactic structure - page 110
Figure 22: A two layer syntactic structure - page 111
Figure 23: The actual structure of the assertion axiom - page 111
Figure 24: A three layer structure - page 112
Figure 25: The proof of identity in Lukasiewicz’s single axiom system - page 117
Figure 26: An example proof net - page 118
Figure 27: A setup for varying languages - page 124
Figure 28: The arithmetic complexity hierarchy - page 147
Figure 29: A window on a single syntactic structure - page 189

Figure 30: Three inter-dependent syntactic structures in separate windows - page 190
Figure 31: The hierarchy of generator classes - page 191
Figure 32: Three patterns - after [194] - page 204
Figure 33: The same patterns with some suggested inclusions - page 205
Figure 34: A diagram of the relationship between relevant types of complexity - page 217
Figure 35: Form and meaning distinction - page 218
Figure 36: Using a genetic population to model an economic agent - page 220

List of Tables

Table 1: Some properties ascribed to the model syntax and semantics - page 38

Table 2: Formulas up to size 6 sorted by size - page 193

Table 3: Formulas up to size 6 sorted by number of distinct variables - page 194

Table 4: Formulas up to size 6 sorted by maximum depth - page 195

Table 5: Formulas up to size 6 sorted by breadth (number of variables) - page 196

Table 6: Formulas up to size 6 sorted by cylomatic number of their minimal graph - page 197

Table 7: Rankings of formula by different methods (bold lines group equally ranked formulas) - page 198

Abstract

This thesis analyses the conception and measurement of complexity and then applies it to some aspects of formal languages.

It starts with a review of the philosophy of modelling. It continues by considering some simple examples to establish intuitions about the common use of ‘complexity’ and goes on to examine what complexity can usefully be attributed to as a property. It argues that it most useful as an attribute of the specification of a model. Some unsatisfactory accounts of complexity are discussed as motivation for the definition of complexity that is then suggested. Some other accounts of complexity are shown to be special cases of the one suggested here.

This approach is then applied to formal languages. A set of properties of analytic complexity are set-out. The set of measures which satisfy these properties is formally investigated. The cyclomatic number of a representation of expressions is put forward to model analytic complexity. In order to analyse shifts in complexity a formal device called syntactic structures is defined. This consists of layers of syntaxes, each with its own production rules which generate the contents of that layer. Each syntactic structure can use substitutions from lower such structures, so that collections of such structures can form hierarchies.

These approaches to are then applied to axiomatic and proof theoretic aspects of logic. Some potential methods of simplification are suggested. Finally some remarks are made about the philosophical applications of this approach.

The appendices include a survey of measures of complexity in the literature; a brief description of a software tool written to explore syntactic structures, two relevant papers on the application of these ideas to scientific modelling and economics, and an extensive bibliography.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

Notes of copyright and the ownership of intellectual property rights

(1) Copyright in the text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged with the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copy made. Further copies (by any process) of copies made in accordance with such instructions may not be made without permission (in writing) of the Author.

(2) The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of Department of the department of Philosophy.