

1 Introduction

1.1 Background

The use of “complexity” as a label for an object of study is fairly recent. Before the middle of this century it appeared to be merely an antonym of “simplicity” (where “simplicity” here means *that property that guides the rational choice between competing theories that are equally supported by evidence* - see section 6.5 on page 129). This sort of “simplicity” follows in a philosophical tradition that is traced back to William of Occam’s famous razor¹. Although there has been some implicit interest with complexity amongst logicians since the 1930s² associated with limiting *size*, the first direct discussion of “complexity” that I have found is in Weaver’s paper in 1948: “Science and Complexity” [464].

In the last 10 years there has been a tremendous amount of interest in “complex systems” of various kinds³ and consequently there has been some interest in “complexity” *per se*. Unfortunately⁴, this has caused an overloading of the term so that frequently it seems only to tag a work as inhabiting an intellectually desirable area⁵. Thus “complexity” appears in the titles of books and articles which, on inspection, are only tenuously related to it. In many of these works it is sufficient for them that they are considering situations where previous analytical techniques fail⁶ or merely that some process of elaboration is involved.

If an author talks about a *complexity measure* then usually more justification is made for the term's use. Many of these measures, however, have a post-hoc flavour; a device is invented or required that has something vaguely to do with complexity and, for the want of a better term, is given that label. Some other, more serious attempts, seem skewed by a prior formal requirement.

1. “Entities are not to be multiplied beyond necessity” – previously invoked by Durand de Saint-Pourcain.

2. See section 5.6.1 on page 113.

3. For an overview of this see [168].

4. Unfortunately for its analytic use that is, the transfer of such terms to the public domain has other social uses [455].

5. This is the essence of Horgan’s critique in [233].

6. Many publications that can be grouped under the “Sciences of Complexity” banner fall into this category.

Some serious attempts at tackling the subject have been made, notably in the mid '70s in general systems theory and recently in the study of chaotic processes. The former were more generally applicable but seemed to peter out due to a lack of consensus and the limited number of practical results gained. The later is still in its infancy but is naturally concerned with the specific problems of modelling chaotic physical processes.

References to complexity are now spread over a wide range of fields, including: biology, mathematics, physics, computation, logic, economics, software design, philosophy, general systems, management science, psychology and linguistics, although the two fields of chaotic process and computational complexity account for many of them.

1.2 The Style of Approach

Firstly, I will focus my discussion on what complexity *is* rather than what might *cause* it. It would seem vain to attempt an investigation of such causes *in general*; if there were but a few general causes then we might well have made more direct progress in tackling the problem of complexity where it arises. Having said that, making progress towards pinning the concept of complexity down would be a first step towards a more coherent and cumulative study into its causes in particular contexts⁷.

Secondly, within this framework I will be approaching the subject from a pragmatic perspective. What I will be looking for are verbal and formal models of complexity that are generally applicable and useful. A useful model is, trivially, one which helps you achieve your goals. In this case it will imply models of complexity which are clear, manipulable (both formally and informally), applicable in a wide range of circumstances and reflect the motivations listed below. I will not go much further into the meaning of utility, as this would go beyond the scope of this thesis. This pragmatic perspective must be borne in mind when interpreting the discussion on the attribution of complexity in section 3.3 on page 47 and its definition in section 4 on page 72.

Such a pragmatic modelling approach seems to fall somewhat between prescription and description. A useful model may be considered prescriptive if you subscribe to the same goals and framework as that of the model. Further than that I do not wish to prescribe usage of “complexity”, except to point out some of the less helpful consequences of its unthoughtful usage. It may be considered as descriptive because it can

7. For an example of such a call for the formal study of complexity see Casti in [89].

be used as a point of reference to relate different examples in an illuminating way. It thus may be used as a descriptive *framework*.

Thirdly, I will apply these models. Although I will consider several examples and consider the philosophical applications, the main target of this study is the application of such models to formal syntactic systems, since this is where I see the numeric quantification of complexity providing the most insight.

1.3 Motivation

The motivation for this study is as follows:

1. *To lay bare assumptions and relativisations involved in the usage of the term.*

As will become clear, the term is frequently used in a vague way, making many assumptions concerning the aims, language and foundation of the subject matter being discussed. These assumptions are sometimes not justified. An example where the vagueness of the term can frustrate progress is in the study of evolution where the issues of whether complexity increases with evolution and how this may happen are debated with cross purposes (see section 6.6 on page 130). I have only found one instance where it was suggested that the rigorous definition of complexity might be counter-productive [260].

2. *To allow a meaningful comparison between different formulations of complexity across different fields of study.*

At the moment, apart from systems theory, models of complexity tend to be formulated with quite specific purposes in mind⁸. The result of this is that there are only vague and unformalised connections between such models developed in different fields. This must impede progress and mean that there is needless repetition. A common explanatory framework would make clear the similarities and differences that exist between such formulations across subject boundaries.

8. This is particularly clear in physics, in the development of models of complexity applicable to chaotic processes, e.g. [457]

3. *To lay the foundations for formalisations of complexity, in different circumstances.*

A formalisation is often more useful for the development of ideas if its components have a clear meaning. These philosophical investigations can provide some of this framework.

4. *To aid the formulation of deeper insights into possible causes of complexity.*

Once one starts to develop models within a coherent framework, it is often easier to build upon previous results and relate informal discussions on the subject given a concrete referential framework that such a model would provide.

5. *To allow the development of systematic approaches to simplification.*

Simplification is the pot of gold at the end of the complexity rainbow. Systematic approaches to simplification (both human and automatic), would be of immense use in academic study, almost irrespective of how marginally it was. Such techniques would be applicable to both the object and method of study, as well as highlighting the cases where simplification is not possible.

This thesis aims to make some progress towards these goals.

1.4 Style of Presentation

Given the above motivation, and that my approach to the problem of defining, formalising and using the concept of complexity will be pragmatic in nature, I am concerned that this thesis should not be only a survey of existing models of complexity⁹. So I will relegate a survey of such papers to Appendix 1, which will summarise the main approaches and give the relevant references. Thus, frequently, when discussing these approaches, I will not swamp the main text with citations but refer to the appropriate section of this appendix, where a more systematic account of each idea can be given and the full references given. In this way I hope to improve the readability of the main text, while retaining its scope.

I have also relegated a lot of the proof details, formalisation, and two papers summarising some of philosophical applications of this approach to complexity to the appendices. Thus in this thesis the appendices contain much of the content of this thesis,

9. If existing measures were more established and demonstrably useful my approach would be different.

but I felt that this arrangement made the central thrust of the thesis more focused and accessible.

1.5 Outline of the Thesis

Section 2 reviews some of the thought on models and modelling in the philosophy of science and the machine learning communities. It then analyses some of the components of the modelling apparatus on a model-theoretic basis, and establishes some terminology that will be used in this thesis. Then it uses this analysis to categorize three ways that different models are related. Finally it briefly state the philosophical position from which the thesis is written.

Then section 3 on page 44 starts with some examples in order to focus on the properties of complexity. It argues that complexity is only usefully attributable to model descriptions relative to a language of expression. It also considers and argues against a number of other basic approaches.

In section 4 on page 72, I present my approach, discuss it and the terms I mention. I also give some examples and relate this approach to some other approaches.

In section 5 on page 86, I apply this approach to formal languages. Here I focus upon what one might mean by “analytic complexity” and what properties one might expect of it. I then present a way of structuring formal languages to enable the analysis of complexity in such languages, which I then apply to axiom complexity and proof complexity. I finish this section by considering some approaches to systematic simplification.

In the last section (section 6 on page 126), I consider some philosophical applications of my approach and conclude in section 7 on page 134.

There are several appendices: section 8 on page 136 is an overview of the main approaches to the formulation of complexity in different fields, section 9 on page 164 holds the details of some of the longer proofs; section 10 on page 182 gives the detailed formalisation for layered syntactic structures; section 11 on page 188 describes a computer-base tool I wrote to help explore the properties of syntactic structures; section 12 on page 193 lists tables with rankings of the complexity of logical formulas according to different measures; section 13 on page 199 is a paper on the philosophical application of complexity to scientific modelling, section 14 on page 210 is an unravelling of various

Syntactic Measures of Complexity

conceptions of complexity in economics and section 15 on page 223 comprises the references.