Positive Philosophy and The Automation of Reason

Roger Bishop Jones

September 10, 2012
# Contents

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>6</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 The Project .......................... 1
1.2 Themes ................................... 2
   1.2.1 The Organon .......................... 2
   1.2.2 Historical Threads ................. 2
   1.2.3 Foundationalism ..................... 3
1.3 By Chapter ......................... 3
1.4 How to Read this Book ............. 5

## 2 The Project

2.1 Leibniz, Characteristic and Calculus ................................... 7
2.2 Carnap’s Programme ......................... 8
2.3 Information Technology ....................... 10
   2.3.1 Alan Turing .......................... 10
   2.3.2 The Science of Computing ............ 11
   2.3.3 Automation and Intelligence .......... 12
   2.3.4 Some Distinctions ................... 12
2.4 The Philosophy .......................... 12

## 3 Fundamental Dichotomies

3.1 Hume’s Fork .............................. 13
   3.1.1 Relations of Ideas ................... 14
   3.1.2 Matters of Fact ...................... 14
   3.1.3 The Place of The Fork in Hume’s Philosophy .......... 15
3.2 A Contemporary Perspective ............. 15
3.3 Before Hume .............................. 15
   3.3.1 The Pre-Socratics ................... 16
   3.3.2 Two Kinds of Stability ............. 17
   3.3.3 Plato’s Theory of Ideals ............ 17
   3.3.4 Aristotle’s Logic and Metaphysics .......... 17
   3.3.5 Rationalist Philosophy ............... 17
   3.3.6 British Empiricism Before Hume .......... 17
   3.3.7 Leibniz .............................. 17
3.4 After Hume ............................. 19
   3.4.1 A Broad Sketch of the Development .......... 20
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Kant</td>
<td>20</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Bolzano</td>
<td>20</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Frege</td>
<td>20</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Russell</td>
<td>21</td>
</tr>
<tr>
<td>3.4.6</td>
<td>Wittgenstein</td>
<td>21</td>
</tr>
<tr>
<td>3.4.7</td>
<td>Tarski</td>
<td>21</td>
</tr>
<tr>
<td>3.4.8</td>
<td>Carnap</td>
<td>21</td>
</tr>
<tr>
<td>3.4.9</td>
<td>Quine</td>
<td>21</td>
</tr>
<tr>
<td>3.4.10</td>
<td>Kripke</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>Analyticity and Analysis</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract Logical Analysis</td>
<td>22</td>
</tr>
<tr>
<td>4.2</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>4.3</td>
<td>The Scope of Analytic Truth</td>
<td>23</td>
</tr>
<tr>
<td>4.4</td>
<td>Analytic Philosophy</td>
<td>24</td>
</tr>
<tr>
<td>4.5</td>
<td>Analyticity in Science</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>Computation and Deduction</td>
<td>26</td>
</tr>
<tr>
<td>5.1</td>
<td>Before the Greeks</td>
<td>26</td>
</tr>
<tr>
<td>5.2</td>
<td>Sketch</td>
<td>27</td>
</tr>
<tr>
<td>5.3</td>
<td>The Status of Proof</td>
<td>28</td>
</tr>
<tr>
<td>5.4</td>
<td>Incompleteness and Recursion Theory</td>
<td>28</td>
</tr>
<tr>
<td>5.5</td>
<td>Computing Machinery and Proof</td>
<td>29</td>
</tr>
<tr>
<td>5.6</td>
<td>Sound Computation as Proof</td>
<td>30</td>
</tr>
<tr>
<td>5.7</td>
<td>Oracles and the Terminator</td>
<td>30</td>
</tr>
<tr>
<td>5.8</td>
<td>Self Modifying Procedures</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>Rigour, Scepticism and Positivism</td>
<td>32</td>
</tr>
<tr>
<td>6.1</td>
<td>Systematic Skepticism</td>
<td>33</td>
</tr>
<tr>
<td>6.2</td>
<td>positivism</td>
<td>34</td>
</tr>
<tr>
<td>6.3</td>
<td>Rudolf Carnap</td>
<td>34</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Tolerance, Pluralism, Metaphysics</td>
<td>34</td>
</tr>
<tr>
<td>6.4</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>Epistemic Retreat</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>Language Planning</td>
<td>38</td>
</tr>
<tr>
<td>8.1</td>
<td>Languages, Notations and Representations</td>
<td>38</td>
</tr>
<tr>
<td>8.2</td>
<td>Universalism and Pluralism</td>
<td>39</td>
</tr>
<tr>
<td>8.3</td>
<td>On the Need for Synthetic Propositions</td>
<td>40</td>
</tr>
<tr>
<td>8.4</td>
<td>Logical Foundation Systems</td>
<td>40</td>
</tr>
<tr>
<td>8.5</td>
<td>Interactive Theorem Provers</td>
<td>41</td>
</tr>
<tr>
<td>8.6</td>
<td>Theory Hierarchy as Knowledge Base</td>
<td>41</td>
</tr>
<tr>
<td>9</td>
<td>The Architecture of Knowledge</td>
<td>43</td>
</tr>
<tr>
<td>9.1</td>
<td>Requirements from Leibniz and Carnap</td>
<td>43</td>
</tr>
<tr>
<td>9.2</td>
<td>Epistemic Retreat</td>
<td>44</td>
</tr>
<tr>
<td>10</td>
<td>Metaphysical Positivism</td>
<td>45</td>
</tr>
<tr>
<td>10.1</td>
<td>First Base</td>
<td>45</td>
</tr>
</tbody>
</table>
CONTENTS

10.2 By Comparison with Logical Positivism ................................................................. 46
10.3 Principal Features .................................................................................................. 47

11 Digging Deeper ........................................................................................................ 48
  11.1 Foundations for Knowledge .................................................................................. 48
  11.2 Logical Foundations ............................................................................................. 48
  11.3 Empirical Foundations ........................................................................................ 51

Bibliography .................................................................................................................. 52

List of Tables ................................................................................................................ 53

List of Positions .......................................................................................................... 55

Index ............................................................................................................................. 55
Preface

[I really don’t have much clue what to do with this preface. The plan for the preface is as follows. Every now and then I will just throw it away and write another. Hopefully, sometime, I will write one that I like, and then keep it.]

This is the first of two volumes in which I hope to give a first account of a philosophical point of view which I call Positive Philosophy. In this preface I describe the scope of the enterprise and the way in which it is divided across the two volumes, by reference to the divisions in the philosophy of Aristotle.

To a first approximation the first volume is concerned with theoretical philosophy, the second with practical philosophy. Of these, theoretical philosophy is concerned with knowledge, with logic, mathematics and empirical science, and practical philosophy with action, with ethics, politics and economics.

These two categories, the theoretical and the practical, appear in Aristotle, but are not exhaustive of his work. Two other divisions are the organon, which consists of Aristotle’s writings on or related to logic, conceived rather more broadly than today, and productive science, which involve arts and crafts, or more generally, how to make or do things. In this two-fold division of positive philosophy, the first volume, theoretical philosophy, includes also logic. In addition to this augmented theoretical philosophy, and the second volume on practical philosophy there is an important element of what Aristotle might have thought productive philosophy, concerned with a kind of production which I think of primarily as engineering, albeit primarily software engineering.
Chapter 1

Introduction

1.1 The Project

Three centuries ago the philosopher, logician, mathematician, scientist and engineer Gottfried Wilhelm Leibniz conceived, as a young man, a grand project, so far ahead of its time that no-one has yet come close to realizing it. The principal elements his project were:

- A universal formal language in which all scientific knowledge might be expressed.
- An encyclopedia of science formalized in that language.
- A calculus or algorithm for answering question expressed in that language.
- Mechanical calculators able to perform the necessary computations.

He devoted considerable energy to progressing these ideas. But his ideas were pie in the sky, there was no hope of success.

Since then many important developments have taken place in philosophy, logic, mathematics and information technology, and most of the now known impediments to the realization of something close to his dream have now been surmounted.

Today it is natural to think that computers solve the mechanical side of the project, and that the rest of the project is just software of various kinds. Though Leibniz did work on the hardware, in his day mechanical calculators, his main interest was in what we would now call software, the algorithms and the data upon which they operated, and his approach to this was philosophical and logical.

This is at once an information technology project and philosophical research. To engineer a “cognitive agent”, to build something which has or can acquire knowledge, and can reason from that knowledge to solve problems, can only be done in the context of a suitable philosophical framework. At its most abstract, architectural design for cognitive artifacts is philosophy.

The kind of philosophy required to underpin such a project is not piecemeal philosophical analysis, it is a systematic philosophical synthesis. It is the aim of this book to undertake such a synthesis, in such a way that its relationship with the engineering of a certain kind of intelligent artifact is made clear. The engineering enterprise I call here “the project”. The philosophy is intended to underpin that project and to constitute its earliest most abstract stages.

It is therefore intended to present by stages together,

- a conception of an engineering project one of whose goals is the automation of engineering design (the aims of the project),
- some relatively abstract ideas on how that project can be organized and implemented (some architectural design)
- a philosophical framework in the context of which the project aims and the proposed architecture can be described and in which the reasons for believing that the architecture might achieve the aims can be articulated.

In this introduction I shall survey the structure of the exposition and sketch the principal themes.

It is in the nature of this project that it combines materials from and seeks to interest devotees of diverse disciplines. The various themes discussed demand varied background, much of which I cannot hope to supply. Though the philosophical the logical and the technological aspects are intimately intertwined, it is hoped that readers with a particular competence and interest in
one aspect will find most of the material of interest to him intelligible without a complete mastery of the other aspects of the presentation.

1.2 Themes

1.2.1 The Organon

The collection of those works of Aristotle which concerned logic is known as the organon. This word comes from the Greek word for a tool.

It has this name because logic was conceived by Aristotle and by later scholars as providing a tool, rather than as a purely academic pursuit. The principal application of that tool was to be demonstrative science, the derivation of necessary truths in the various sciences from the first principles of those sciences.

The emphasis on logic as a tool might well have been unimportant through most of the history of logic, since until recently Aristotle’s organon has dominated the field. However, though well intentioned, Aristotle’s formal logic is inadequate for any non-trivial scientific application, and the study of logic has remained the province of Philosophers.

With the advent of modern logic, beginning in the second half of the nineteenth century with mathematicians and philosophers such as Boole, Frege and Pierce, there was spawned a new discipline of mathematics, mathematical logic, which was primarily meta-theoretical in character. Academic interest in logic now spans multiple disciplines, of which the most important are philosophy, mathematics and computer science. There are various ways in which these disciplines make use of logic as a tool, but its use as a tool in the manner envisaged by Aristotle, as the means whereby conclusions are drawn from various first principles, is rare.

This book belongs to a line of philosophical works (notably, Aristotle, Leibniz and Carnap) of which the aim is to contribute to the means for the application of formal deductive methods in the establishment and application of knowledge. This work is subordinate to that purpose.

The realization of that objective depends upon a philosophical context which cannot be taken for granted, or even taken from a shelf and dusted off. Over the last half-century philosophy has moved in other directions, and in order to do so has undermined the philosophical basis for the most recent manifestation of the project in the philosophy of Rudolf Carnap. For a resumption of the project, new philosophical foundations are required.

1.2.2 Historical Threads

The philosophical innovation required in support of the project consists substantially in the re-establishment of ideas which have fallen into disrepute. This includes specific ideas such as the analytic/synthetic dichotomy, and entire philosophical perspectives or systems such as positivism. Their reestablishment, at least as viable alternatives to received opinion, will not be realized by a detailed refutation of the arguments which displaced them.

In logic over the last 150 years and in Computing over the last 50 years there have been very many new developments of a kind which one might expect to give philosophers pause for thought. But new ideas in philosophy itself are very rare, most of the twists and turns in philosophical fashion are revivals of old ideas in new clothes. It is in the nature of philosophical progress that it often appears through millennia of debate as ideas are proposed, developed, disputed, rejected and perhaps ignored for a while before rising again re-engineered for a new philosophical climate.

If we seek afresh to understand and to advance such ideas, tracing their development through history may be helpful in getting or in conveying an understanding of their contemporary manifestation.

Several historical threads serve I hope to illuminate aspects of the present proposal. The philosophers whose ideas are touched upon in these sketches have spent a lifetime developing the ideas. The purpose of the sketches is to make the ideas presented here clearer by connecting them with their historical roots.

The first of these historical threads sketches the analytic/synthetic dichotomy and a variety of related dichotomies and concepts. This is closely connected with the development of ideas of logical truth, and of truth conditional semantics.

Alongside such questions relating to the establishment of meaning, there is the question of truth and how it may be established or evaluated. Doubts about our ability to establish truths are at their most severe in scepticism which part of the historical background to and continuous with positivism, which combines elements of scepticism with a strict attitude towards rigour in science. Our positivism departs from its predecessors
substantially in ways which can be illuminated by considera-
tion of those sceptical roots, and which make it natural to think of as a graduated positive scepticism. This is closely connected with the question of rigour, and with the tension between rigour and progress, in science, mathematics and philosophy.

Though there appears to be at any point in time a trade-off between rigour and productivity (which is perhaps easiest to see in mathematics), at times advances are made which allow both to advance at once, and a new balance to be achieved. This happened in the nineteenth century for mathematics, a period of consolidation in which standards of rigour in mathematical analysis were transformed. The last stages in this transformation involved the invention of modern methods in logic, and established the possibility of the formal derivation of mathematics. The greater precision in locating the most abstract subject matters of mathematics, through the agency of axiomatic set theory, resulting in achievements of high standards of rigour which were sustained through a century of continuous mathematical innovation. These higher standards depended only peripherally on the new mathematical logic. Axiomatic set theory provided sufficient additional clarity to the definition of mathematical concepts, that standards of rigour were able to advance without the adoption of formal derivation as the standard for mathematical proof. It sufficed for a mathematical to convince his peers that a formal derivation would be possible.

In the second half of the century, the application of digital computers to the support of formal notations and deductive systems has created a new domain in which for the first time formal notations and languages are extensively used. Stored program computers demand and support the use of such formal languages. Communicating unambiguously with computers becomes a prime motivation for the use of formal notations, which provide not only the motivation but also the kinds of support which facilitate the use of formal languages.

Epistemology is approached in what might perhaps be described as an instrumental manner. Two aspects of epistemology receive no attention. The first is the meaning of the word “know”, and hence for some philosophers the question of what knowledge is, is not considered. Second is any aspect of how knowledge is acquired or applied which is peculiarly human.

Instead of considering what knowledge is we are concerned with how knowledge might be represented and applied.

The epistemology is foundational both in relation to logical and empirical knowledge. The distinction between these two, the logical and empirical, is a keystone of the project and the philosophy. The project realizes an general analytic method a principal feature of which is the systematic and thorough separation of these two kinds of knowledge. Logical knowledge (under a broad conception of logical truth as analyticity) is represented formally and established by deductively sound methods. Empirical knowledge is represented formally using abstract logical models. The relationship between abstract formal theories and the systems which they model is ultimately informal.

1.3 By Chapter

The Project Chapter 2.

The first stage of this, in Chapter 2, is to give the next stage of detail in describing the project, the aims and the architecture. This is presented by starting with an account of Leibniz’s project, tracing the history of the ideas from the 17th through to the 21st century, and then offering a revised conception of such a project for the 21st.

The philosophical side of this history ends in debacle. The last major philosophical proponent of a significant fragment of the Leibnizian enterprise was Rudolf Carnap in whose philosophy the formalization of science had a central place. In mid 20th century, fundamental philosophical ideas with a key place in Carnap’s philosophy were repudiated by W.V.Quine. In particular, the analytic/synthetic distinction, subject to continuous critique by Quine since his first exposure to Carnap’s philosophy, was given a full bleded and uncompromising repudiation in Quine’s influential “Two Dogmas of Empiricism”. Though not aligning himself with Quine on the analytic/synthetic dichotomy Saul Kripke then teased apart the triumvirate of concepts which had been identified by Carnap (analyticity, ne-
cessity and the *a priori*), allowing Kripke to inaugurate a new kind of metaphysics, and set analytic philosophy on a new track fundamentally at odds with the philosophy of Carnap. Thus Kripke contributed to the subsequent widely held view that the philosophy of Rudolf Carnap (sometimes known as *Logical Positivism*) had been decisively refuted as a result of technical advances by two of the most highly competent and respected philosopher-logicians of the period.

**Fundamental Dichotomies** Chapter 3.

The philosophical framework I offer here is closer to the philosophy of Carnap than to that of any other philosopher, and it is therefore necessary in Chapter 3 to repair some of the damage done to this point of view. Chapter 3 considers the status of these three dichotomies, the distinctions between analytic and synthetic, between necessity and contingency, and that between the *a priori* and the *a posteriori*. Though not always with this vocabulary, similar distinctions have been talked of throughout almost the entire history of western philosophy. Through this history one can see both a gradual refinement of these concepts and also reversals. I therefore lightly trace this history showing how Carnap’s understanding of these dichotomies was reached, and then review and respond to some of the modern criticisms which are still held by many to be decisive against Carnap. In this chapter I argue that the refutation of Carnap on these matters is one more demonstration of the contrast between the rigour of mathematics and that of philosophy. An illustration, I might say, of the *irrationality* of philosophy.

**Analyticity and Analysis** In Chapter 4 I then take the concept of analyticity as established, and the question of its significance is examined in greater detail. In philosophy the twentieth century was called the *age of analysis*, and the principal kind of philosophy progressed by academics was called *analytic philosophy*. In the philosophy of Rudolf Carnap and the logical positivists the connection between the concept of analyticity and analytic philosophy was simple. Insofar as philosophy was concerned with establishing the truth of propositions (in the manner in which mathematicians establish results by proving theorems, or science establishes physical laws by observation and experiment) the results of the kind of philosophical analysis envisaged by Carnap would be analytic, though in his hands such results play a secondary role to the articulation of methods and the definition of languages or concepts suitable for science. For none of the many other conceptions of philosophical analysis which appear in the 20th century was there such a simple connection between analyticity and analysis. In this chapter we look at the relevance of analytic truth to various kinds of analysis, both philosophical and scientific. This is done using a kind of philosophical thought experiment. Suppose that we had an oracle (man or machine) which could tell us of any conjecture whether or not it was an analytic truth? What impact would that have on the various kinds of analysis under consideration?

**Computation and Deduction** Chapter 5.

With the concept of analyticity and hence of *deductive soundness* in place it is time to further refine my characterization of the project, as being concerned, firstly, with *deductively sound computation* and ultimately with useful approximations to the *terminator*. These terms are explained in Chapter 5, and a variety of contemporary research trends are compared with the research thus envisaged. Since the beginning of mathematical logic many different conceptions of have evolved of proof and its relation to computation. The approach envisaged here is clarified in the context of a general discussion of these different conceptions of proof.

**Rigour Skepticism And Positivism** Chapter 6.

Carnap’s philosophy had been intended to provide a way forward for philosophy to the achievement of standards of rigour comparable to those of mathematics (an ideal which has been held by many philosophers over the last 2500 years), but his program had been defeated, illustrating just those defects that he sought to remedy. To reinstate this idea I sketch another historical thread in Chapter 6. This is a history of rigour in mathematics and in philosophy. It is a history also of those philosophers who have perceived the rational deficit in philosophy, or in the search for knowledge more generally, of the *sceptics* of ancient Greece, and of the more modern tradition of *positivism* consisting of a kind of mitigated or constructive scepticism in which high standards of rigour are articulated for both *a priori* and *a posteriori* sciences.

**Epistemic Retreat** Chapter 7.

The epistemologically conservative aspects of positivism give rise in this philosophy and architecture to
the notion of “epistemic retreat”. This involves an admission of general doubt, but the acceptance of degrees of doubt, and hence a partial ordering of conjectures indicating in a relative way how well they have been established, and what level of confidence they may be viewed.

Language Planning  Chapter 8.

Carnap’s pluralism, a willingness to accept the use of any well-defined language on a pragmatic basis, gives rise to the problem of “language planning”, addressing for example the problems arising from the use of multiple different languages for different aspects of the same problem. In Chapter 8 the architecture is further developed by addressing these matters.

Carnap built on the purely mathematical foundational ideas of Frege and Russell, but sought to apply the new logical methods to the empirical sciences. He believed that this required innovation on his part to admit languages suitable for talking about the material world rather than purely about mathematics, and in making this transition he also moved from a purely universalistic conception of logic (in which one language sufficed) to a pluralistic conception of the language of science. His contribution to this pluralistic world was by way of meta-theory, he spoke about how languages might be defined in their syntax, semantics and proof rules, taking this to be a proper philosophical contribution to the methodological advancement of science. The proliferation of languages thus envisaged would demand some kind of activity which he called “language planning”, but did little on.

The project I envisage embraces the pluralism of Carnap, and therefore depends upon an architecture which admits multiple languages and permits large scale applications involving more than one language. This connects with other initiatives in computing, notably the idea of an “Extensible Markup Language” (XML) and the many related ideas which have built up around it, including the idea of a “semantic web”.

So far as applications to empirical science are concerned the project we outline does not adopt the approach of Carnap. Instead we envisage that the project provides or empirical science (and ultimately of various engineering enterprises which depend upon it) by regarding these as working exclusively with abstract models of the physical world, and take the connection between such models and the concrete world to be beyond the scope of these formal methods.

Instead of asserting the truth of an empirical theory we instead evaluate the contexts in which it provides a useful model of aspects of the real world, and evaluate the model in different application domains in terms of reliability and fidelity or precision. Carnap’s notion of language planning (on which he himself said little), is one domain in which our project might best be compared with the W3C Semantic Web initiative.

The Architecture of Knowledge  In chapter 9 we now table an architectural proposal, in the form of a set of key requirements and a set of architectural features intended to realize those requirements, and a rationale for the belief that they do indeed realize them.

Metaphysical Positivism  With the architectural proposal in place we return in chapter 10 to the philosophy.

This first involves gathering together a coherent and rounded philosophy sufficient to underpin the proposed architecture, an important element of this is foundational. The second part concerns the methods supported by the architecture and their scope of applicability.

Digging Deeper  Metaphysical Positivism does not answer all philosophical problems, but it does influence what might be considered a worthwhile philosophical problem for further investigation. Chapter 11 looks at some of these.

Climbing Higher  Where does this architecture take us, why should it be implemented?

Chapter ?? is concerned with the outer reaches of the significance of the project, beyond the confines of philosophy or of academia.

1.4 How to Read this Book

Far be it for me to say how you, reader should proceed. However, here are some ideas, and some observations on how I have tried to write it which might be helpful.

I rarely myself read a book linearly from cover to cover. The book nominally addresses a very broad range of potential readers, many of whom will be interested in only some aspects of its subject matter. In writing it I have therefore tried to make it possible for readers to reach those parts which matter most to them.
without having to struggle through too great a jungle of detail which might seem to them peripheral.

To this end I have tried to begin and end each chapter with summary material which for some readers might suffice, and to include references as specific as possible in the text to prior materials upon which an understanding may depend.

I have felt it desirable, in order to make as clear as possible the ideas which I present, to make use of stories about the history of various aspects of the subject matters. Often the work of philosophers who have spent a lifetime producing an important body of original work will be spoken of in a few sentences which cannot be a fair account of their work, even in some special corner.

In order to avoid misrepresentation I have used wherever possible the device of enunciating a position which, whether it was ever held by any philosopher or not, is useful in making a point. One or more philosophers may be named as having inspired this position, without going into a detailed examination (which I am rarely best equipped to undertake) of how closely it does correspond to the positions they in fact held.

This is a method not unrelated to that of the philosopher Saul Kripke in his examination of certain ideas suggested to him by the writings of Ludwig Wittgenstein. The method may be adopted of connecting a philosophical problem which is thought to be of interest in its own right with the history of the subject. Alternatively, for those whose interest is primarily historical and exegetical, the process of rational reconstruction may begin in this way, with a definite model of some aspect of a philosopher’s work, which may cast light by evaluation of its similarities and differences with the textual sources, and which may be refined in the light of such comparisons into progressively more complex and subtle models less readily seen as diverging from the target of analysis.

In this work it is the former motivation which concerns us exclusively. The second, as a kind of analysis is of interest from a meta-theoretic point of view, but is not here practiced in anger.

Too rigorous an attempt to distill historical illustrations into hypothetical positions not directly attributed would however be unduly cumbersome. This mode of presentation is reserved for the most important and substantial points, and much background is presented as if historical fact, but should nevertheless be thought of in a similar manner, as so simple an account as could at best be true in spirit, only to be dissipated on closer inspection.
Chapter 2

The Project

It is my aim here to devise a philosophical framework suitable to underpin a particular enterprise, the project. This might sound like a narrowly scoped philosophy, but “the project” is broad enough to encompass the whole of mathematics, science and engineering when undertaken by deductively sound methods. The project is, in brief, the automation of deductive reason and its applications, undertaken in a particular manner which I hope to make clear. The relationship between that project and the philosophical ideas which I offer as a supporting context is complex. A full and detailed account of the project depends upon the philosophy, but the project plays an important role in the articulation of the philosophy, they are hand-in-glove.

In this section I make a first attempt at describing the project trying to minimize reference to the underpinning philosophy. In subsequent chapters the philosophical issues are entered into in earnest, and in this way it is hoped that aspects of the project are made more clear.

The scope of the phrase “automation of deductive reason” is intended to include intelligent reasoning, but not to encompass much of the aims or methods of that field known as artificial or machine intelligence.

Here and in subsequent sections explanations I use try to make the ideas clearer partly by talking about their historical evolution. In this section the connection with some of the pioneers and visionaries who have contributed to the formalization and automation of reason in ways particularly relevant to our project are drawn out. The principle figures in this are Gottfried Wilhelm Leibniz and Rudolf Carnap, as exemplars of the philosophers with similar projects, and Alan Turing is discussed for a contrast with the kind of automation of reason I have in mind here.

2.1 Leibniz, Characteristic and Calculus

Leibniz was a rationalist, a philosopher who was particularly concerned with what can be established by reason rather than observation. The distinction between the two he recognized in the dictum:

“There are two kinds of truths, truths of reason and truths of fact.”

Despite recognizing this distinction, he perceived that all these truths might be codified formally, so that the distinction between truths and falsehoods could be determined by computation.

His vision was universalistic in several respects. He envisaged a universal characteristic, which was to be a formal logical language in which all knowledge might be codified, and a calculus ratiocinato, a method of calculating the truth value of any sentence in his universal characteristic.

These two ideas provide the first prototype for this work.

Leibniz did not imagine that the truth of scientific hypotheses could be established in this way, rather that the corpus of scientific knowledge could be encoded in the universal characteristic, the calculus then settling questions in the context of that knowledge. In consequence, Leibniz also promoted the collaborative development of an encyclopedia, of scientific academies and journals. Leibniz perhaps also guessed that the computations involved might be non-trivial, and contributed also to the development of the information technology of his age, calculating machines.

Leibniz in these ideas was at least three centuries ahead of his time. The logic he knew was essentially
CHAPTER 2. THE PROJECT

that of Aristotle, and fell short of what is needed for the formalization of science by a wide margin. The necessary innovations in logic did not appear for another two hundred years. Beyond the adequacy of modern formal languages to express the entire corpus of scientific knowledge, there lie known limitations in the extent to which formal deduction can capture truth in these systems. Mechanical decision procedures are known not to exist, so for these reasons we know that no calculus could be relied upon to settle all well formulated problems in the universal characteristic.

These limitations are perhaps of only marginal significance for applications in science and engineering, but even within these limitations there are serious problems of computational intractability. For many or most problems whose answer is in principle computable, it will not be accomplished in a reasonable timescale or through the deployment of available resources. Brute computation will therefore not suffice in any but the most simple cases, and something closer to the workings of human intelligence would be needed, to come close to an effective implementation of Leibniz’s calculus.

The technology of calculating machines was even further from sufficiency. It remains moot whether today’s information technology is up to the task. Beyond the adequacy of the hardware, there are problems with the software, potentially even less tractable.

Leibniz’s ideas in these matters had little influence in his own time or for centuries following, not just because they were so far removed from what was then realizable, but also because they were not even published. His largest impact pressed matters in the opposite direction, and came from his independent invention of the differential and integral calculi, and particular for the notation which he devised for this epoch making mathematical innovation. The epoch thus spawned was two centuries of rapid development of mathematics applicable in science and mathematics. These enormous advantages took place despite the reduction consequent on the introduction of infinitesimal quantities. Right at the beginning of this period the philosopher Berkeley criticized these new developments, but it was two hundred years before mathematicians began to take the problems of rigour seriously and the rigorization of analysis began, ultimately leading to the new developments in logic which made it sufficient for Leibniz’s project.

The rigorization of analysis was accomplished first by the displacement of infinitesimals in a precise notion of limit and by construction of real numbers from natural numbers, effectively reducing analysis to arithmetic. This done, the next challenge taken up by Kettle’s Frege was to show that arithmetic was simply a matter of logic, requiring no special metaphysical insights. In rising to this challenge Frege, inspired in part by the example of Leibniz, re-invented logic in the form of his Begriffsschrift or concept notation, which provided the basis not merely for the logicisation of arithmetic, but also potentially for Leibniz’s universal characteristic.

Frege’s project was more narrowly scoped, but even the formalization of arithmetic was an arduous undertaking, into which Frege had invested decades of industry when Bertrand Russell pointed out to him the inconsistency of his basic principles. At about the time when Russell and Whitehead were completing their own assault on the logicisation of mathematics (a substantial hurdle in which was ensuring against the kind of inconsistency found in Frege’s system) the young Rudolf Carnap was attending as an undergraduate Frege’s lectures on his new logic and absorbing much of Frege’s attitude to rigour in deductive reasoning and its general applicability.

2.2 Carnap’s Programme

Carnap was first apprised of modern logic as an undergraduate by Frege’s lectures on his ‘Begriffsschrift” [Fre67, VH67]. At this stage in his development Carnap clearly showed an interest both in philosophy and in mathematics and science. He was also strongly inclined towards the precision of language which he found in Frege’s logic which he contrasts with the teaching of logic in philosophy. Within the sciences, he preferred physics because of its greater conceptual clarity. He also perceived the distinction between ethical, evaluative, emotive and metaphysical language and scientific doctrine.

From this position as an undergraduate just before the great war he moves forward in a period of about seven years to the point at which we can see the formulation of the central ideas which motivated his work through the rest of his life. The most significant new influence in forming these ideas came from Bertrand Russell. Carnap became acquainted with Principia Mathematica[WR13] and began to prefer its notation to that of Frege. He also begins to feel that a concept is not
clearly understood until he can see how to express it in symbolic language.

At the end of 1921 Carnap read Russell’s *Our Knowledge of the External World*[Rus21] and was inspired by Russell’s characterization of “logico-analytic method” in philosophy. It is at this point that Carnap self-consciously devotes himself to this philosophical method and begins intensive reading of Russell’s writings on the theory of knowledge and the methodology of science.

Russell, in his work on *Principia Mathematica* had undertaken with Whitehead a task similar in character to Frege’s project. However, Frege’s focus had been on the logicisation of mathematics. Russell had a broader conception of the applicability of the methods, and advocated a kind of philosophy in which such methods were used exclusively. Carnap took up this challenge enthusiastically.

Carnap’s philosophical programme therefore involved first the idea that philosophy in general should be *analytic* in the specific sense that its methods and results are logical, and should be obtained by the new logical methods pioneered by Frege and Russell. Secondly Carnap conceived of the role of the philosopher as being primarily concerned with facilitating the progress of empirical science. The kind of facilitation he had in mind was analogous to the innovations in mathematics undertaken by Frege and Russell, the establishment of new languages in which scientific laws could be precisely enunciated and the surrounding theoretical and methodological framework for the formalization of science.

The work of Frege and Russell had been *universalistic* in the sense that they sought a single new language in which the whole of mathematics could be developed. At this stage Carnap’s thinking was along similar lines, but since he was trying to carry forward the ideas into empirical science, it did not seem feasible to stick with the same purely logical systems. Carnap’s aims at this point may therefore be thought similar to Leibniz’s desire for a *universal* characteristic in which all of science could be formalized. Later Carnap became more pluralistic, but the idea of using formal logic to encode scientific knowledge and of formalizing deductive reasoning about science places his enterprise within the scope of Leibniz’s.

Like Leibniz, Carnap’s ideas were ahead of their time. He was working in the context of modern symbolic logic, which is (I shall argue), no longer shackled by the limitations of Aristotelian syllogistic. But strict formality in language and proof is arduous, and good mechanical support a prerequisite to widespread adoption of such methods. *Principia Mathematica* formalized a significant part of mathematics, and was very influential during the formative years of the new academic discipline of *mathematical logic*. These new developments did have a significant impact on Mathematics, which was made more rigorous through the systematic use of axiomatic set theory. But mathematicians did not follow the example of Russell and Whitehead by taking up strictly formal notations or formal proofs. Not even in mathematical logic did this happen, for mathematical logic became a meta-theoretic discipline in which formal systems were studied rather than applied. Proofs in mathematical logic may often be concerned with formal systems, but they are not themselves formal.

Not only was formalization rejected by mathematics and science, it was soon to be rejected, together with Carnap’s entire philosophical outlook, by philosophy. The first major assault on Carnap’s position was by Quine, who had been sniping at some of the core doctrines, notably the concept of analyticity, ever since his first exposure to Carnap’s philosophy as a Harvard postgraduate visiting Carnap in Vienna. Quine’s critique modulated into outright repudiation of the central tenets of Carnap’s philosophy in his “Two Dogmas of Empiricism”. It was not long before an attack came on another flank from Saul Kripke, who dismantled the identification of analyticity and necessity which were inter-definable in Carnap’s conceptual scheme and thereby invented a new kind of metaphysics.

The idea of Russell and Carnap, that philosophy should consist of logical analysis and of Carnap that it should also be conducted formally both disappeared. Symbolic or mathematical logic does play a substantial role in analytic philosophy, but it is the meta-theoretic techniques of mathematical logic which are used in philosophy. There is no general adoption of formality as a way of achieving rigour in philosophy, no general recognition that there is a deficit in rigour might requires such a remedy.

Another impediment to the realization of Leibniz’s project was, during Carnap’s lifetime, beginning to be dissolved. The digital electronic computer was invented and the first steps toward using this technology for the automation of deduction were in progress. The torch was about to pass from philosophy, not to mathematical logic, but to Computer Science.
2.3 Information Technology

These two philosophers strove to realize ambitions which were soon to be made more feasible by the invention of the stored program digital computer, and with it another academic discipline, Computer Science, which over the second half of the 20th century would conduct an enormous amount of research relevant to both of their projects. The first home of logic, Philosophy, and the new discipline of mathematical logic, treated logic in a meta-theoretically. They took formal logic primarily as something to be studied rather than directly used (by contrast with the earlier work by Frege and Russell and Whitehead in which large parts of mathematics were formally derived). Computer Science, as well as conducting theoretical and meta-theoretical investigations relevant to computing, found reason to undertake proofs formally, with the assistance of their computing machinery.

For these reasons, the primary academic locus of research relevant to the ambitions of Leibniz and Carnap moved from Philosophy to Computer Science. However, the idea of using computers for science was more predominantly pursued by purely computational rather than formally logical methods. I hope to make this distinction clearer, to make the case that logical methods deserve to be given special consideration, to give a philosophical context in which it make sense to do that on a substantial scale, and to consider some aspects of how this might be progressed.

2.3.1 Alan Turing

Alan Turing was a mathematical logician, computational engineer and a visionary thinker whose ideas on artificial intelligence have been influential on research in the automation of reasoning.

For our present purposes it is primarily his writing on artificial intelligence which is of interest, as encompassing objectives even broader than our own, but Turing was an important figure in a milestone in our understanding of computation which was reached during the 1930s.

One of the tenets of the philosopher and mathematician David Hilbert in the early part of the 20th century was that all definite mathematical problems are susceptible of solution. This thesis was held to be equivalent to the effective decidability of validity in first order predicate logic, which was called the entscheidungsproblem. To make the problem precise enough to be given a definitive mathematical answer it was necessary to make precise the notion of effective calculability. This resulted in several different logicians putting forward different notations in which arbitrary algorithms (methods of solution) could be expressed. Church came up with concise notation for functional abstraction known as “the lambda calculus”, Kleene with a system for defining numerical functions by recursion, Emil Post came up with a system based on transforming strings using a set of “productions” which defined transformations on the strings. All these systems provided ways of defining effectively computable functions over the natural numbers, but the one which looked most like a description of a computing machine was that invented by Alan Turing, since called the Turing machine [Tur36].

These different ways of describing effective computational processes all turned out to be similarly expressive, and this remarkable discovery underpinned the idea that they did indeed all capture the notion of effective calculability. The entscheidungsproblem having by these means been made more definite, Church and Turing independently solved the problem by exhibiting problems which were provably unsolvable (by any effective procedure).

This result limits what could possibly be achieved by Leibniz’s calculus ratiocinator, as did Gödel’s previous result on the incompleteness of arithmetic. It seems from Leibniz’s description that he probably did imagine that any definite question within the scope of the scientific knowledge codified in the lingua characteristica would be answerable using his calculus and a sufficiently rapid calculator. The Gödel and Church-Turing results show that this cannot be the case. I will look a little closer later at these and other limiting factors and consider their significance for “the project”.

Turing also wrote on Artificial Intelligence, and his conception of Artificial Intelligence provides an alternative ideal for the automation of reason which is not predicated on machines achieving the impossible. His seminal paper in this area was Computing Machinery and Intelligence[Tur50]. Turing’s essential idea was that human intelligence, though very differently implemented, is not fundamentally distinct from or more capable than the kind of information processing which could be undertaken by a sufficiently complex and powerful Turing machine. In addressing the question whether a machine can think, Turing described an “imitation game” to make this question more precise, this later became known as the Turing test. The
Turing test involves human beings interacting with a person and a machine through similar interfaces which conceals which of the two was involved, i.e. using a keyboard to converse with someone or something in a distant or private place. If the observer cannot tell the difference between the man and the machine when he interacts with it in this way, then perhaps we should be ready to acknowledge that the machine thinks, and hence exhibits some important aspects of human intelligence.

2.3.2 The Science of Computing

During the first half of the 20th century the technology of computation had moved rapidly. The Turing machine provided a very simple abstract prototype for an important transition, from special purpose to general purpose computational engines. It anticipated the general purpose stored program digital computer, and these relatively complex calculating machines were just becoming technically feasible. They first appeared in the 1940s using technologies such as electromagnetic relays, and the electronic vacuum tube. Soon the transistor emerged, and the technology of computation began a long period of progressive miniaturization yielding ever smaller and faster building blocks for digital computers.

With the invention of the digital computer came the new academic discipline of Computer Science, and enthusiasm for formal systems and the automation of reason passed from Philosophy and Mathematical Logic (itself just an infant) to Computer Science. Computer scientists were interested in formal languages because they allowed algorithms to be described precisely for execution by a computer. The complexity of these algorithms rapidly increased and correctness became an issue. How could one be sure that the instructions for achieving some computational objective would realize their intended aim? Since formally defined algorithms could be construed as mathematical entities, it seems that certainty might best be realized by a mathematical (i.e. a deductive) proof. Unfortunately these proofs turned out themselves to be even more complex than the programs whose correctness they were intended to show, and the questions then arose how one could facilitate the construction of these proofs and how one could be sure that the proofs were correct? The computer itself provided one answer to these questions. The computer could assist in the construction of the proofs, and since it is in the nature of formal proofs that their correctness can be checked mechanically, they could also check the proofs. In this way, that part of Computer Science which was particularly interested in ensuring that computer programs are correct was drawn into the use of formal notations, formal deductive systems and the automation of reason.

Meanwhile another branch of Computer Science developed which sought to program computers so that instead of merely doing large scale drudgery, they exhibited some kind of intelligence. This was the field of Artificial Intelligence, and later the related interdisciplinary enterprise of Cognitive Science. Within these disciplines a variety of approaches to the problem were pursued, but despite this variety a single important cleavage was captured by the contrast between “neats” and “scruffs”. The project around which this book is constructed may be thought of leading towards an extremely “neat” AI, so this distinction can be helpful in initial sketches of the project.

The distinction between neats and scruffs may be seen through the contrast between “connectionist” AI and methods based on theorem proving. The connectionist approach to AI approximates the human brain by simulation of a neural network using highly parallel processing. The idea is to devise a network which is capable of learning and then to teach it the required skills. If such a system achieved a competence in a complex deductive science it would do so indirectly. Such competence would be an application of its general intelligence, not a source or means of attaining that intelligence.

“neats”, on the other hand, think more like Leibniz. For them the intelligent capabilities are build on a foundation which includes a formal language for representing propositional knowledge and a deductive inferential capability. This does not necessarily mean that they are aimed at different problem domains to the “scruffs”. Neats may seek systems capable of ordinary language discourse and common sense reasoning, but still approach this on logical foundations which would be unfamiliar to most people.

The neat/scruffy distinction is concerned with means rather than ends. There is a related distinction which concerns ends rather than means. This is the distinction between taking the aim of a research programme to be the replication of some aspect (or the whole of) human intelligence, of whether some other characterization, independent of the methods or competence of
human beings, is give of the objective of the research.

2.3.3 Automation and Intelligence

We arrive then at my first characterization of the project which this philosophical dissertation seeks to underpin.

The aim of the project is a comprehensive formalization of the deductive aspects of philosophy, mathematics, empirical science and engineering design.

2.3.4 Some Distinctions

Developments since Leibniz and Carnap give us better insights into the extent to which their projects are realizable. The diversity of related research which has been undertaken enables us to discriminate many different similar kinds of project and to identify more precisely the kind of projects which Leibniz and Carnap might have entertained if they were working today.

2.4 The Philosophy

Here are the bare bones of a philosophical framework to underpin “the project”.

The philosophy is positivistic, insofar as it delineates a particular approach to empirical science and its application, which it is suggested may in some ways be preferable to extant methods. In this I am not being prescriptive, I do not assert that science should be conducted in this way.

The framework is, in the first instance, conceptual. This consists to a large extent in the choice of particular meanings for terms which have a long history of shifting usage. Some of the most fundamental of these come in pairs as dichotomies, i.e. disjoint concepts which together exhaust some significant larger domain. The most fundamental of these is the distinction between logical and empirical truths. Simple though it might appear to be, it is a source of endless philosophical controversy, and it will therefore be necessary to consider in detail the evolution of this distinction over its entire history and come to as precise an understanding of this fundamental cleavage as possible.

Such fundamental concepts are not mere accidents of language, and are one element of the philosophical framework which we may think constitute a kind of metaphysics.

1 On this my disagreement with Carnap lies only in the scope of applicability of the term metaphysics. It does not constitute metaphysics under either of the two principle categories which constitute metaphysics for Carnap. I take great care to make the meaning of these concepts definite and clear, so meaninglessness, I hope, is avoided. Nor does this kind of metaphysics constitute, in our framework, a priori justifications about synthetic claims.
[In this chapter, as in all chapters which are primarily historical, it is particularly important not to slip into too purely chronological an account. All the life is in the particular themes which the narrative is intended to illuminate, and it is the development of these themes which must be at all times in the foreground. I do not know how to achieve this. Part of the difficulty in producing the story is that the detailed content of these themes has to be well understood in order to find a good way of presenting them, but one feels that the required level of understanding can only be extracted from the history.]

We saw in the last chapter how the most recent successor to Leibniz’s project, in the philosophy of Rudolf Carnap, was derailed by the rejection of some of the most fundamental tenets of Carnap’s philosophy.

It is conceivable that the project could be revived on a different basis, for neither Quine nor Kripke, despite their devastating critiques, actually abandoned the idea of deductive reason or the use of formal deductive systems. If I was myself convinced of the soundness of those criticisms then I might follow that course, but I am not. It is best instead to answer them.

In this chapter I therefore focus on the concept of analyticity and its relation to those of necessity and of the \textit{a priori}. Before addressing specifically the criticisms of Carnap’s position by Quine and Kripke, I propose to sketch the history of the evolution of these concepts over the past two and a half millennia. I will present this as falling in two principal stages, first the development of predecessors of these fundamental concepts and then the subsequent refinement of our understanding of the concept. The point of transition I suggest, is with Hume, though this is more an expository device than a dogmatic claim. The idea that there exists any such definite point of transition is tenuous, but the supposition helps to give structure to the narrative.

It is with Hume, I suggest, that we find a first account of the dichotomy which is not easily seen to be in some respects defective. This is perhaps as much to do with Hume not seeing himself as defining the distinction, but rather as drawing attention to it and placing it in a central place in his philosophy.

Before entering into the history I will sketch the technical content of the narrative, to provide a framework in which the evolving detail can be placed. This is done using the ideas of Hume.

### 3.1 Hume’s Fork

David Hume was a philosopher of the Scottish Enlightenment. The enlightenment was a period of ascendency in the place of reason in the discussion of human affairs, when science had secured its independence from the authority of church and state and had a new confidence in its powers derived substantially from the successes of Newtonian physics.

Hume looked upon the philosophical writings of his contemporaries and found in them two principal kinds, an “easy” kind which appealed to the sentiments of the reader, and a “hard” kind which trawled deeper and appealed to reason. This latter kind, “commonly called” metaphysics, was preferred by Hume, but found nevertheless, by him, to be lacking, infested with religious fears and prejudices. Hume’s feelings about these aspects of philosophy were not vague misgivings. He had a specific epistemological criterion which he saw these philosophical doctrines as violating.

Hume’s project involves an enquiry into the nature of human reason for the purpose of eliminating those parts of metaphysics which go beyond the limits of knowledge, and establishing a new metaphysics on a solid
foundation limited to those matters which fall within the scope of human understanding.

David Hume wrote his philosophical magnum opus, *A Treatise on Human Nature* [Hum39] as a young man. He was disappointed to find his work largely ignored and otherwise misunderstood, and thought perhaps that his presentation had been at fault. To improve matters he wrote a much shorter work more tightly focussed on the core messages which he thought of greatest importance, which he called *An Enquiry into Human Understanding* [Hum48].

In a central place both logically and physically in this more concise account of his philosophy he says:

“All the objects of human reason or enquiry may naturally be divided into two kinds, to wit, Relations of Ideas, and Matters of Fact.”

We shall see that Hume is here identifying a single dichotomy which corresponds to all three of the distinctions which here concern us. In his next two paragraphs he expands in turn on the kinds he has thus introduced.

### 3.1.1 Relations of Ideas

“Of the first kind are the sciences of Geometry, Algebra, and Arithmetic; and in short, every affirmation which is either intuitively or demonstratively certain. That the square of the hypotenuse is equal to the square of the two sides, is a proposition which expresses a relation between these figures. That three times five is equal to the half of thirty, expresses a relation between these numbers. Propositions of this kind are discoverable by the mere operation of thought, without dependence on what is anywhere existent in the universe. Though there never were a circle or triangle in nature, the truths demonstrated by Euclid would for ever retain their certainty and evidence.”

Hume is distinctive here among empiricist philosophers in having a broad conception of the *a priori* (though he does not use that term here), allowing notable the whole of mathematics. In this he may be contrasted for example with Locke who allowed only certain rather trivial logical truths to be knowable *a priori*. Nevertheless, Hume’s conception of the *a priori* remains narrow by comparison with the rationalists, and in particular, as Hume will later emphasize, excludes metaphysics.

### 3.1.2 Matters of Fact

“Matters of fact, which are the second objects of human reason, are not ascertained in the same manner; nor is our evidence of their truth, however great, of a like nature with the foregoing. The contrary of every matter of fact is still possible; because it can never imply a contradiction, and is conceived by the mind with the same facility and distinctness, as if ever so conformable to reality. That the sun will not rise to-morrow is no less intelligible a proposition, and implies no more contradiction than the affirmation, that it will rise. We should in vain, therefore, attempt to demonstrate its falsehood. Were it demonstratively false, it would imply a contradiction, and could never be distinctly conceived by the mind.”

The evolution of following three dichotomies are the subject matter, though we will find other related dichotomies which feature in the history.

The terms which I will use to speak of them, in this chapter are:

- analytic/synthetic
- necessary/contingent
- *a priori/a posteriori*

As I shall use these terms these are divisions of different kinds of entity, by different means.

The first is a division of sentences, understood in sufficient context to have a definite meaning, and is a division dependent upon that meaning.

The second is a division of *propositions*, which may be understood for present purposes as *meanings* of sentences in context. The division is made according to whether the proposition expressed must under all circumstances have the same truth value, or whether its truth value varies according to circumstance. In this we are concerned with two particular notions of necessity, those of logical and of metaphysical necessity, the latter being sometimes taken to be broader than the former. A part of the role of Hume’s fork in positivist philosophy is to banish metaphysical necessity insofar as this goes beyond logical necessity.
The third is for our purposes also a division of propositions, on a different basis. It concerns the status of claims or of supposed knowledge of propositions. It is expected that such a claim must in some way be justified if we are to accept it, and that the kind of justification required depends upon the proposition to be justified. The justification is a priori if it makes no reference to observations about the state of the world.

3.1.3 The Place of The Fork in Hume’s Philosophy

The mere statement of the fork (which we shall see, is not original in Hume) is of lesser significance than the role which it plays in Hume’s philosophy, which serves to clarify the distinction at stake and draw out its significance.

Hume’s philosophy, like Descartes’ comes in two parts of which the first is sceptical in character, and the second constructive. In both cases the sceptical part clears the ground for a new approach to philosophy which is then adopted in the constructive phase.

For our present purposes we are concerned principally with the first sceptical phase of Hume’s philosophy, because of the delineation of the scope of deductive reason, and hence of the analytic/synthetic dichotomy which is found in Hume’s sceptical arguments. This delineation is baldly stated in Hume’s first description of the distinction between “relations between ideas” and “matters of fact”, for there Hume tells us that no matter of fact is demonstrable.

This bald statement would by itself have little persuasive force if it were not followed up with more detail, even though ultimately this detail does not so much underpin the distinction as depend upon it.

Hume’s further discussion begins with the consideration of what matters of fact can be known ‘beyond the present testimony of our senses or the records of our memory’. The inference beyond this immediate data is invariable causal, we infer from the sensory impressions or memories to the supposed causes of those impressions. But these are not logical inferences, causal necessity is for Hume no necessity at all (even less the inference from effect to cause). Hume’s central thesis that matters of fact are not demonstrable is in this way reduced first to the logical independence of cause and effect, and then to the distinction between deductive (and hence sound) inference and inductive inference (whereby we infer causal regularities and their consequences).

Given that Hume considers all inferences from senses to be based on induction, and sees no validity in causal inference, it follows that from information provided directly to us by the senses nothing further can be deduced which is not simply a restatement, selection or summary of the information itself. Further enlightenment from this sceptical doctrine is primarily the application of this principle to various kinds of knowledge. In the process Hume does a certain amount of

3.2 A Contemporary Perspective

In tracing the history of the analytic/synthetic and related distinctions I hope to show how their various historical manifestations relate, and to present a perspective from which the various developments may be evaluated perhaps as plain progress, perhaps as advances in certain respects, perhaps as involving or constituting regress.

Such evaluations can only be made from a particular point of view, and in this section that point of view is outlined, and related, in the first instance to Hume’s fork.

3.3 Before Hume

The distinctive place of Hume in the history of “the fork” which bears his name is not a mark of his priority in making the distinction, but rather of an important point in its development, a clear identification of exactly the right distinction (I now suggest) of the central place which he gave it in his philosophy, and of the evidence he gives (in what is are now regarded as Hume’s various sceptical doctrines) of a sound intuitive grasp of the scope and limitations of logical truth and deductive inference.

To underpin the significance of Hume’s distinction, and to see that, notwithstanding its apparent simplicity, it is by no means simple to arrive at, I now look at some of the most important stages in the earlier development of the ideas.

The philosophers I will consider here are Socrates, Plato, Aristotle, Descartes, Locke and Leibniz.

Most of the elements which we find in Hume’s fork can be found in the earliest of these philosophers.
The relevant ideas which these philosophers discussed include Plato’s world of ideals and of appearances, the notions of essential and accidental predication, and of necessary and contingent truth, and Aristotle’s logic and the idea of demonstrative truth. Leibniz is important for his conception of the scope of logic and the possibility of arithmetisation and mechanisation. His intended method for these depends upon a conceptual atomism which was to exert significant influence on the philosophy of Russell and of the Early Wittgenstein, and thus indirectly upon Carnap’s conception of logical truth.

The distinction between the logical and the empirical influenced two major tendencies in modern philosophy, namely rationalism and empiricism of which Leibniz represented the first and Hume the second. The distinctive feature of these tendencies has been respectively the overestimation and underestimation of the scope of deduction, championed by metaphysically and scientifically inclined philosophers (or philosophically inclined scientists). The dialogue between these has therefore served to refine the distinction between the logical and the empirical. These two tendencies may be seen to have been anticipated by Plato and Aristotle, of whom Plato seems like the rationalist, and Aristotle is closer to empiricism. Curiously, the connections with Leibniz and Hume are crossed over; it is Plato the rationalist who seems to provide the clearer anticipation of Hume’s fork, and it is Aristotle the less rationalistic of the two who provides the logic on which Leibniz’s ideas were built.

Hume’s principle description of his fork is in terms of subject matter, the distinction between knowledge of ideas and knowledge of the world. In Plato we see something quite similar, he distinguishes between the eternally stable world of Platonic ideals, of which we can have certain knowledge obtained by reason, and a world of appearances, in constant flux, and of whose fleeting nature we can at best have tentative opinion based on the unreliable testimony of our senses. Not only does the distinction of subject matter match, but the key characteristics of these domains, that in the one we have solid precise, reliable knowledge, and in the other, nothing of the kind, are agreed between them.

For an understanding of the philosophy of Socrates and Plato it may be helpful to consider first the pre-Socratic philosophers.

3.3.1 The Pre-Socratics

The history of the analytic/synthetic dichotomy is connected with that of deductive inference, which is generally held to have begun with Greek mathematics.

It might be argued that the ability to undertake elementary deductions is an essential part of competence in a descriptive language. For to know the meaning of concepts, one must also know at least the more obvious cases of conceptual inclusion. One cannot be said to know the meaning of the word “mammal” if one does not know that all mammals are animals, and hence that any general characteristic of animals is possessed by mammals. The ability to draw inferences purely based on an understanding of language does not however entail the capability to discriminate between such deductive inferences and inferences which are based not merely on meanings but also on supposed facts, so no clear grasp of what inferences are deductive need be involved. Furthermore, competence in drawing conclusions may by entirely unwitting, there may be not awareness of the idea of inference.

It is in Ionia in the 6th century BC that mathematics was transformed into a theoretical discipline, a key feature of which is the practice of proving general mathematical laws (rather than simply recording them for general use). At this time philosophy, the love of knowledge, encompassed mathematics and science as well as what we might now call philosophy. The use of reason in mathematics was very successful, and Greek mathematics grew into a substantial body of rigorously derived, knowledge of which much is known to us through the thirteen books of the Elements of Euclid (c300 BC). In Euclid’s geometry, developed deductively using an explicitly documented deductive axiomatic method, Greek mathematics achieved a standard of rigour which was not to be surpassed for two thousand years.

The success of deduction in mathematics was not reflected in other aspects of the work of the pre-Socratic philosophers. It is distinctive of Greek philosophy, beginning with the Ionians, that religion and myth was not a dominant influence, and that philosophers sought knowledge by rational means, rather than deferring to any kind of authority. Ionian philosophers, often described as cosmologists, sought unifying principles, as a way of understanding the diversity of the world around them. Different philosophers adopted different hypothesis about the ultimate constituents of the uni-

---

1 A part of classical Greece now belonging to coastal Anatolia, itself a part of Turkey.
verse. Thales held that the world originated from water, Anaximenes from air. Empedocles held that all matter was formed from air, water, earth and fire. By contrast with mathematics, Greek philosophy became a great diversity of irreconcilable doctrines.

The philosophical search for unity and stability underlying apparent diversity and flux yielded no universally accepted result. This is nowhere more conspicuous than in the philosophies of Heraclitus and Parmenides, the first asserting an external flux in which nothing remained stable and the second that nothing changes. The failure of reason to resolve differences in this domain of enquiry was underlined by the arguments of Zeno of Elea, who devised many paradoxical arguments to reduce to absurdity the possibility of change, in support of the philosophy of Parmenides. In this way it is made conspicuous that apparently the same method, reason, works very well for mathematics but fails miserably in metaphysics.

3.3.2 Two Kinds of Stability

We can understand the search for the pre-Socratic philosophers for the unity and stability underlying the diversity of the world by analogy either with modern science or with mathematics.

In the case of science we seek physical laws which govern the changes which take place in the world. The world changes, but the laws are immutable. In modern science the preferred way of formulating scientific laws is using mathematics. The scientific hypothesis is that the relevant aspects of the world correspond more or less exactly with the structure of some mathematical model.

The unchanging fundamental truths may therefore be sought either in empirical scientific laws or in mathematics.

These two alternatives connect with tendencies in modern philosophy, empiricism and rationalism, and to aspects of the two great philosophical systems of classical Greece, those of Plato and Aristotle. The systems of Plato and Aristotle in different ways anticipate Hume’s fork, and in the rationalist and empiricist tendencies we can see a dialectic on the scope of deduction through which the distinction evolved towards Hume’s formulation of the fork.

3.3.3 Plato’s Theory of Ideals

Plato sought to place philosophy on as rigorous a footing as mathematics, and he did this with a synthesis of the ideas of Heraclitus and Parmenides which recognized two distinct domains of rational discourse. Mathematics succeeded because it reasoned deductively about abstract ideas.

In Plato Hume’s fork is anticipated as a distinction between two worlds, the world of ideal forms and the world of appearances. Thought of in this way the distinction is about subject matter,

- **Distinction 3.1** the world of forms and world of appearances
- **Distinction 3.2** the a priori (knowledge) and a posteriori (opinion)
- **Distinction 3.3** essential and accidental predication

3.3.4 Aristotle’s Logic and Metaphysics

3.3.5 Rationalist Philosophy

3.3.6 British Empiricism Before Hume

Hume was one of a line of British philosophers who emphasized the role of experience in the acquisition of knowledge. Important figures in this tradition were Bacon, Hobbes and Locke.

3.3.7 Leibniz

Leibniz conceived of the project to which this book is devoted. We are concerned here with just one aspect of his work, which is his contribution to the ideas leading to Hume’s fork.

Through most of the intellectual history right up to the 20th century Aristotle is a dominating influence on thinking about all aspects of Hume’s fork. Some of the influence of Aristotle is negative, his subject-predicate analysis of propositions and his syllogistic conception of logic were both to narrow and adherence to Aristotle’s ideas may have inhibited the development of logic sufficient for our purposes.
Leibniz conceived of the idea that reason might be automated as a young man, and pursued this project for the rest of his life. His insight was specifically about how the Aristotelian syllogism could be automated through arithmetisation, thus anticipating a method which was to become famous in logic when applied by Gödel in his incompleteness results.

Within this context the logical and metaphysical ideas which underpinned and provided technical substance to his project are important.

Like Aristotle, Leibniz did have the concepts of Necessary and Contingent truth, he also had the concepts of \textit{a priori} and \textit{a posteriori} truth which were closely connected, as in Hume. There is also a connection with semantics, similar to that in Aristotle, through their role in \textit{a priori} proof and hence in the establishment of necessary truths. To this is added a precise description of a mechanizable process of analysis whereby necessary truths might be established, which differs greatly in character from any previous work in logic.

Though Leibniz’s work in this area (by contrast with his work on the calculus) exerted little influence in his time, but proved an inspiration for some of the leading figures in the development of logic in the 19th and 20th centuries, including Gottlob Frege and Bertrand Russell. His influence is conspicuous in Russell’s philosophy of logical atomism[Rus18] and Wittgenstein’s \textit{Tractatus Logico-Philosophicus} [Wit61].

Leibniz was a \textit{rationalist} philosopher and his conception of the division which concerns us sometimes sounds radical. He holds for example that, at least for God, all truths are necessary. But he nevertheless does distinguish between necessary and contingent truths and, for us mere mortals the dividing line between these two falls more or less where Hume will put it. Mathematics is necessary, science is contingent. He also closely connects, as will Hume, necessity and \textit{a priori}.

In relation to the analytic/synthetic distinction, Leibniz still uses these concepts as they are used in classical Greece, to describe two different approaches to logical proof, contemporary usage of “synthetic” comes from Kant. Leibniz does have something like a semantic and a proof theoretic characterization of necessary and \textit{a priori} truth. On the proof theoretic side it is of interest not just that Leibniz considers necessary truths to be those susceptible of a priori \textit{proof}, thus implicitly distinguishing the method of verification from the manner of discovery. On the semantic side we have an explanation of necessity and of proof via the analysis of definitions.

Leibniz tells us that \textit{truth} is a matter of conceptual containment, the containment of the concept of the predicate in the concept of the subject. This kind of containment we would now describe as \textit{extensional} and would not accept as a characterization of logical or necessary truth, but rather of truth in general, including contingent truths. There is another kind of conceptual containment which applies only to necessary truths, and that may variously described as \textit{intensional}, \textit{essential} or risking circularity, as necessary or analytic. This is the containment of \textit{meanings} and descriptions of the role of definitions in the process of proof are consistent with the view that logical truths correspond to containment of meanings rather than containment of extensions.

This distinction does not feature in Leibniz, his distinction between necessary and contingent propositions for us mortals is made in two other ways. The first is by appeal to the omniscience of God and the limitations of man. Men are not capable of the kind of complete comprehension of meaning which God exhibits. Propositions which are contingent for us are those whose analysis is beyond the limits of our knowledge or analytic capability, but will nevertheless, if true, be knowable by analysis for God. The second way of making the distinction is through the principle of “sufficient reason”. Logical necessities can be shown to be true by reduction to the law of identity of which self-predication is an instance. The procedure is to analyze the predicate and subject by expanding their definitions and discarding components present in the subject but not in the predicate, until the subject and predicate turn out to be identical. The remaining necessities arise from “the principle of sufficient reason”, which tells us that nothing happens without sufficient reason, and more specifically that the world is the way it is because from the logically possible alternatives available to God, he has chosen the best.

We seem to have here two kinds of regression from the position occupied by Aristotle. We see a reduction in the clarity of the concept of demonstrative truth, by introducing the idea that the definitions of concepts may be beyond our ken, and intelligible only to God. This replaces the distinction between essence and accident, which is closer to the idea that in one case \textit{meaning} suffices, and in the other observation will prove necessary.

Leibniz’s ideas on mechanisation of the syllogism flow from a conceptual atomism. The word “term” is used
in Aristotelian logic for the subject and predicate of a proposition in subject/predicate form, both of which may be thought of in more modern language as expressing concepts. Leibniz’s atomism is the idea that there are simple concepts from which all other concepts are formed in specific ways, together with the retention from Aristotle of the idea that all propositions have subject/predicate form.

In this it is important to distinguish the simple concepts from simple expressions which may be used to express them. Leibniz atomism here is not about syntax, it is about those things which the syntax expresses, the underlying concepts. A complex concept may nevertheless have a simple name, a complex concept is one which is defined in terms of other concepts, a simple concept is one which has no such definition. The atomistic thesis is that there are simple concepts, and that all complex concepts have a definition directly or indirectly in terms of simple concepts.

More specifically, a complex concept can always be analyzed as a finite conjunction of literals (in today's terminology) where a literal is either a simple concept or the negation of a simple concept. Leibniz saw that if every distinct simple concept were represented by a unique prime number, then a conjunction of concepts could be represented by the product of primes. The fundamental theorem of arithmetic tells us that any such a product can be factorized to retrieve the primes corresponding to the simple constituents. He proposed to represent arbitrary complex concepts by a pair of such products, i.e. a pair of numbers, of which the first number is the product of the primes representing the concepts which occur positively in the definition of the complex concept, and the second number is the product of the primes representing the concepts which occur negatively in the complex. The same concept cannot appear both positively and negatively, or else the definition is inconsistent, so these two numbers will be relatively prime (having no common factors). If Leibniz were correct then once the simple concepts have been identified and given unique prime numbers, all complex concepts are representable as pairs of co-prime numbers.

It is then possible to describe (quite simply) calculations which determine the truth of any proposition of the four forms which Aristotle uses in his syllogistic, (when these propositional schemes, which include variables instead of definite concepts in the subject and predicate places, are instantiated with specific concepts).

3.4 After Hume

Hume’s fork abolishes a certain conception of metaphysics, making a difficulty in establishing any other. What it leaves is something that looks rather like science, falling into two parts. One part contains only matters devoid of empirical content, albeit including the whole of mathematics. The other contains opinions of the empirical world, obtained by guesswork based on sensory impressions.

Not many philosophers, or scientists find this a very attractive picture. Much philosophical reaction to this seeks to refute the general scepticism in Hume’s position, but the refutation of his empirical scepticism is more of benefit to science than to philosophy. The significance to philosophy of Hume’s fork is more acute in relation to those special kinds of knowledge, beloved particularly to Plato, which are the truest subject matter of philosophy, and were later to be known as metaphysics. From this point of view positivism, first amply exemplified in Hume, is an anti-philosophical philosophy, in which philosophy gives up its own true ground yielding knowledge to science.

It is understandable that many philosophers will react specifically against this central feature of positivism, and in this section we will consider some of these reactions.

These we consider in three aspects.

1. Firstly there arises in the ongoing dialectic, further refinement of our knowledge of exactly where the fundamental line is drawn.
2. Secondly there are challenges to and reaffirmation of the idea that a single dichotomy is involved.
3. Finally, coupled with the previous two there are various ways of reviving and of dismissing the possibility of some kind of metaphysics.

Alongside these philosophical themes there are two technical problems which are gradually addressed. In order for the dichotomies to be definite it is necessary for languages to have definite meaning. One way to achieve this is to define new languages, and for them, give a definite semantics and deductive system. The ability to do this appears on the scene very late, little more than a century ago. The history both before and
after that point of transition of ideas about semantics and proof, is part of our present concern.

The clarity which I find in Hume’s description of his fork is like a moment of lucidity in life of confusion. Not until Rudolf Carnap do we find a return to and a positive refinement of that distinction.

3.4.1 A Broad Sketch of the Development

It would be easy in a sketch of the subsequent history of Hume’s fork to lose the central issues in the detail of the very considerable developments in logic and philosophy since then.

To help make draw these out the headlines are sketched here before a little more detail is supplied.

Kant supplied the first challenge to Hume’s conception of a single dichotomy. Kant rescued a limited conception of metaphysics by separating the *a priori* from analytic truth (which is contrasted with synthetic truth), and declaring that our knowledge of arithmetic, of space and time were *synthetic a priori*. It is not clear here whether Kant’s conception of analyticity differed materially from Hume’s relations between ideas, or whether the concept was the same but Kant disagreed about its extension.

3.4.2 Kant

The first challenge we consider came from Kant, who was awoken from his dogmatic slumbers by Hume, and was moved to reject the unity of the triple dichotomy.

Kant supplied the first challenge to Hume’s conception of a single dichotomy. Kant rescued a limited conception of metaphysics by separating the *a priori* from analytic truth (which is contrasted with synthetic truth), and declaring that our knowledge of arithmetic, of space and time were *synthetic a priori*. It is not clear here whether Kant’s conception of analyticity differed materially from Hume’s relations between ideas, or whether the concept was the same but Kant disagreed about its extension.

3.4.3 Bolzano

With Bolzano we see a first approach to the definition of concepts relevant to Hume’s fork which advance significantly beyond Aristotle.

The techniques used by Bolzano are a considerable advance, and the conception of logical truth he promulgated is quite close to the idea of logical truth which has dominated mathematical logic ever since. However, this is a narrower concept than the ones we have considered so far, and depends upon the classification of concepts into logical and non-logical.

Though the technical apparatus which Bolzano deploys is a great advance, the concepts he defines with this apparatus take us further away from a notion of logical or analytic truth which is properly complementary to the notion of empirical or synthetic truth.

3.4.4 Frege

Frege’s logical and philosophical work was in part aimed at overturning Kant’s critique of Hume, in particular Kant’s refusal to accept that mathematics is analytic. To achieve this aim Frege invented a new kind of formal logical system of which the two most important features were the abandonment of the Aristotelian emphasis on the subject/predicate form of propositions. Instead of predicates Frege worked more generally with functions, among which predicates are functions whose results are always truth values. Crucially, logical sentences now have arbitrary complexity, and the universal quantifier is introduced to operate over an arbitrarily complex propositional function.

For Frege the notion of logical and analytic truths are separated. Logical truth similar to Bolzano’s in being defined through a separation of concepts into logical and non-logical, but analytic truth in Frege represents the most precise formulation of Hume’s truths of reason to that date.

In the analysis of Frege’s contribution the notion of free logic becomes significant, (though this is not a term he uses). This is so because in his *Begriffsschrift or concept notation* those aspects of deductive logic which do
3.4. AFTER HUME

not involve ontology, or questions of what exits, are satisfactorily addressed. However, when he later comes to apply essentially the same logical system to the development of arithmetic it is necessary to incorporate axioms which suffice to establish an ontology sufficient for mathematics, and at this stage Frege introduces principles which are not logically consistent.

3.4.5 Russell

Russell began his work on mathematical logic rather later than Frege, having been born in 1872, not long before the publication in 1879 of Frege’s Begriffsschrift. Furthermore, because Frege’s work was ignored, Russell did not become aware of it until he was already well progressed with his own approach to the logicisation of mathematics.

Prior to this work Russell had already studied and written his own account of the work of Leibniz, and Russell’s philosophy is significantly influenced by Leibniz. Russell perceived Leibniz as having been handicapped by too closely following the logic or Aristotle, most particularly his retaining the idea that all propositions have subject/predicate form. Russell’s logical ideas build on the work of Pierce and Schröder on the relational calculus. The relational calculus provided a more general form for propositions. A sentence might in effect have multiple subjects of which jointly some predicate is asserted (predicates involving multiple subjects are called relations).

The divergence from Aristotle in this matter was independently progressed by Frege and by Pierce and Schröder.

3.4.6 Wittgenstein

While studying engineering at Manchester, Ludwig Wittgenstein became interested in philosophy through an interest in the logical foundations of mathematics. On advice from Frege, Wittgenstein went to Cambridge to study under Russell at a time when Russell was deeply involved in the completion of Principia Mathematica[WR13].

3.4.7 Tarski

3.4.8 Carnap

3.4.9 Quine

Having studied Russell’s Principia Mathematica even as an undergraduate,

3.4.10 Kripke

The following table gives a chronology.

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
<th>Author</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>585 BC</td>
<td>Thales</td>
<td>Mathematics</td>
<td></td>
</tr>
<tr>
<td>428-348 BC</td>
<td>Plato</td>
<td>The Theory of Forms</td>
<td></td>
</tr>
<tr>
<td>1632-1704</td>
<td>Locke</td>
<td>relations between ideas v. matters of fact</td>
<td></td>
</tr>
<tr>
<td>1646-1716</td>
<td>Leibniz</td>
<td>the synthetic, a priori</td>
<td></td>
</tr>
<tr>
<td>1711-1776</td>
<td>Hume</td>
<td>the method of variations in set theory</td>
<td></td>
</tr>
<tr>
<td>1724-1804</td>
<td>Kant</td>
<td>Begriffsschrift, sinn and bedeutung</td>
<td></td>
</tr>
<tr>
<td>1781-1848</td>
<td>Bolzano</td>
<td>the theory of types</td>
<td></td>
</tr>
<tr>
<td>1845-1918</td>
<td>Cantor</td>
<td>the cumulative hierarchy</td>
<td></td>
</tr>
<tr>
<td>1848-1825</td>
<td>Frege</td>
<td>logical truths as tautologies</td>
<td></td>
</tr>
<tr>
<td>1848-1825</td>
<td>Russell</td>
<td>definition of truth</td>
<td></td>
</tr>
<tr>
<td>1889-1951</td>
<td>Wittgenstein</td>
<td>logical syntax, the method of intensions and extensions against the analytic/synthetic distinction</td>
<td></td>
</tr>
<tr>
<td>1901-1983</td>
<td>Tarski</td>
<td>and modal logics, holism separating analyticity, necessity and a priority via rigid designators</td>
<td></td>
</tr>
<tr>
<td>1891-1970</td>
<td>Carnap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1908-2000</td>
<td>Quine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1940-</td>
<td>Kripke</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Development of the Analytic/Synthetic distinction
Having established the distinction between logical and empirical truths we now turn to a closer consideration of logical truth.

There are two main considerations. One is the scope and relevance of logical truth. This should include a discussion of its place in analytic philosophy, in mathematics, empirical science, and engineering. A principle aim in this discussion is both to re-emphasize, as Hume did, the severe limits to what can be established purely by deduction, and to show that it nevertheless is of the greatest practical significance. The limitations and potential are to be made real by illustrations of a different character to those of Hume but which are connected with the various concerns of my own positive philosophical thinking, theoretical and practical. Specifically in relation to philosophy, some discussion of the nature of philosophical analysis making the distinction between a claim being analytic and a claim being “about language”, i.e. between subject matter and epistemic status.

I would like to introduce here the notion of an analytic oracle, which can then be refined in the next chapter to the FAn oracle.

We have seen that the words “analytic” and “synthetic” acquired in the philosophy of Kant a sense distinct both from their use outside academic philosophy and from previous philosophical and mathematical usage.

In non-philosophical use the related terms analysis and synthesis have diverse application, generally concerned with taking apart or putting together the parts of some complex whole. In the special domain of logical proof, the usage in classical Greece was similar, connoting two methods of proof. An analytic proof proceeded by analysis of the proposition to be proven, ultimately reducing it to principles which can be known without proof. A synthetic proof begins instead with axiomatic principles reasoning forward until the desired theorem is eventually reached. In contemporary automation of reason, these different methods, which we will consider further in a later chapter, are sometimes known as backwards and forwards proof respectively.

Kant introduced a new usage in which analytic and synthetic are applied to propositions (in the Aristotelian sense), classifying them along the lines of Hume’s fork. Kant gave two criteria for this classification, one proof theoretic (concerning the manner in which such a proposition might be established), and the second semantic concerned with meanings. However characterized, analyticity in this sense is closely connected with deductive reason, and it is our purpose in this chapter to relate this precise technical concept with the very general notion of analysis, both in its academic and its more worldly applications.

### 4.1 Abstract Logical Analysis

Though we are concerned here with analysis in general, there is one particular kind of analysis which will have special place. This kind of analysis is closely coupled with the notion of analyticity through the concept of deductive soundness.

The notion of soundness is applicable both to formal logical systems which are supplied both with a semantics (an account of the meanings of the sentences of the language) and a formal notion of derivation or proof, or to particular inferences in any language for which the semantics is well-defined. In the former case the system is sound if the inference rules respect the semantics in such a manner that from true propositions
only true conclusions can be derived. In the latter case, without reference to any particular deductive system we may say that an inference is sound if the premises entail the conclusion. A set of sentences (the premises) entails some other sentence (the conclusion) if under all circumstances whenever the premises are true the conclusion will also be true.

The connection with analyticity is then that all claims about entailments are true if and only if analytic. An alternative statement of the connection is that all propositions derivable from the empty set of premises, the theorems, of a sound deductive system are analytic.

Because of this connection, Rudolf Carnap and the logical positivist held that philosophy, which they thought of as an a priori science consisted of analytic truths. This involves a position on the demarcation of philosophy which we will not adopt here. Instead we simply note the scope of this particular kind of analysis, and of the kind of philosophy which employs it. Philosophy falling within the scope of such methods can now be made as rigorous and reliable as mathematics by the use of modern formal languages and computer software which assists in the construction of formal proofs and automatically checks that the proofs are correct.

Abstract logical analysis consists in the application of deductive reasoning to analysis of arguments, of concepts, of theories or doctrines. The application to arguments may be considered to be primary, and in this case the idea is that in any domain in which reason is thought to be applicable, modern logical methods can be used to improve the rigour of the reasoning. The idea here is much as it was in Plato, it is that the standards of rigour which are generally found in mathematics and generally lacking elsewhere, can be made more widely available by appropriate methods. In particular, by focussing on the concepts involved, by ensuring that we have clear definitions of these concepts, we can reason within the relevant domain rigorously.

Plato, Aristotle, and the many logicians who followed them until very recent times, failed to realize this ideal of logically rigorous reasoning beyond mathematics. The greater difficulty in pinning down non-mathematical concepts may have been a factor here, and it is not until the 19th century that logic was progressed to the point at which the logic itself could be used in making precise definitions from which one can reason with formal rigour.

4.2

Other points of controversy important to this project remain, concerned with the scope, applicability and importance of analytic truth. These are important to us here for two distinct kinds of reason.

The theoretical core of Positive Philosophy, *Metaphysical Positivism*, is primarily an analytic philosophy, and it is essential in presenting a conception of analytic philosophy to address some of the reasons why a purely analytic philosophy might be thought to be of narrow scope and limited value.

The project we are considering is of broader scope, just as were the projects of Leibniz and Carnap, encompassing not merely an approach to philosophical analysis, but also important and substantial parts of mathematics, empirical science, engineering and other activities in which deductive reasoning might play a significant role.

There is a circle to be squared here. There is one perspective from which the entire enterprise is without content, and represents a preoccupation with academic trivia, and another diametrically opposed perspective in which it is of the greatest practical significance and may be thought to warrant substantial and energetic prosecution.

4.3 The Scope of Analytic Truth

I have presented a history of the evolution of the concept of analyticity and related concepts, and have adopted in Metaphysical Positivism a conception which is similar to that described by Hume “relations between ideas”. A good recent account of this concept may be found in the writings of Rudolf Carnap, but more importantly analyticity is a characteristic of the theorems of a large class of modern tools for constructing and reasoning in formal logical theories. Further sharpening remains of interest from a philosophical point of view, and will be discussed later, but for practical purposes, even its relevance to applications demanding the very highest standards of rigour, the concept and our technologies for checking when it applies are sufficient.

It remains a matter of controversy how significant analyticity, analytic truths, and methods in which they figure prominently are or might be.

Analytic truths may be thought of as falling into two principal groups.

The first group consists of true claims in languages
whose subject matter is entirely “abstract”. For this we understand abstract entities as mere ideas, about which we reason by offering a “definition” of the relevant domain of entities and then reasoning logically from the definition to conclusions about that domain which will be true by definition. This group encompasses the whole of mathematics. It also encompasses reasoning about abstract models of any domain whatever, whether or not one considers the model to be “mathematical”, so long as the model is well-defined and the method is deductive.

This is a Platonic conception of the scope of deductive knowledge, and can be broadened very broadly in the direction which Aristotle took, to yield a body of analytic truths which might be thought to be about the material world rather than purely concerned with abstract entities. There are several ways in which this can be done in the context of modern logic without embracing the complexities of Aristotelian metaphysics.

Carnap’s approach is to adopt formal languages whose subject matter is the material world, to define the semantics of the languages by giving the truth conditions, and to define analyticity in terms or such a truth conditional semantics. The analytic truths are then those sentences in such languages whose truth conditions are invariably satisfied.

### 4.4 Analytic Philosophy

The term analytic philosophy was first applied in the 20th century to a kind of philosophy which began at the turn of the century as Bertrand Russell and G.E.Moore. These two men, though sharing the idea that philosophy should be in some sense analytic, had quite different conceptions of the kind of analysis involved, and their differences remained significant as methods of analysis evolved throughout the 20th century.

Russell conception of analysis was shaped by the new methods in logic in the development of which he played an important part. He believed that failures of rigour in philosophical reasoning resulted in some cases from imperfections in ordinary language, and could be avoided if a special ideal language were adopted along the lines of the Theory of Types which he devised with A.N. Whitehead for the formalization of mathematics in Principia Mathematica. He did not advocate or practice the adoption of such formality in philosophical reasoning, but did advocate the adoption of similar logical methods. An example the kind of method he envisaged is the use of logical constructions. By such means Russell advocated a “scientific” philosophy in which logical methods transformed philosophy into a rigorous deductive discipline progressively advancing in a manner similar to that of mathematics and science, rather a perpetual sequence of conflicting theories and a lack of solid progress. We may recall Plato’s prescription here, in which the subject matter of philosophy and the only place where true knowledge might be attained is in the world of ideal forms. Even though Russell was an empiricist, his conception of philosophy is as a deductive discipline.

On the other hand Moore’s conception of analysis concerned natural languages and consisted in the clarification of concepts in those languages, yielding illumination consistent with common sense. There is a connection here with Socratic method, but neither Socrates nor Plato conferred such authority upon ordinary language and common sense. Socrates sought the true nature of concepts such as justice and virtue, and though he believed that ordinary men were in some sense possessed of these true concepts, they nevertheless might not correctly grasp them until they are induced by a Socratic dialogue to properly recall this knowledge.

Is it the case that analytic philosophy in the 20th century was narrower in its scope than philosophical tradition from which it grew, and if so is this a necessary consequence of the conception of philosophy as analytic, or of some particular ideas of what kind of analysis was at stake.

Two mid century conceptions of the nature of analytic philosophy illustrate the issue, those of Rudolf Carnap, exhibited in his Philosophy of Logical Syntax and the subsequent developments of his ideas, and those of the “linguistic philosophy”, exemplified by J.L.Austin, which prevailed briefly in post war Oxford.

A central thesis of Carnap’s philosophy of logical syntax was that philosophy consists of logical analysis and yields results insofar as they are established truths, which are themselves analytic. This apparent narrowing of scope is confirmed by explicit exclusion of fields such as ethics, not only from the domain of analytic philosophy, but from scientific discourse altogether (bearing in mind that this kind of philosophy is itself conceived of as scientific in character, though not empirical). The severity of the apparent narrowing here is mitigated by the predominance in the actual philosophy of Rudolf Carnap of work which is methodological, including work which ensues in some proposal for the use
4.5 Analyticity in Science

of language. Though philosophy has relinquished any claim to offer factual enlightenment about the world, it may nevertheless make material contributions to the advancement of such knowledge by contributions to the methods of science.

Linguistic philosophy in its most extreme form conceives of philosophy as the analysis of natural language.

4.5 Analyticity in Science
Chapter 5

Computation and Deduction

This chapter will include some of the history of computation and logic, showing the relationship between these two topics, showing the changes to them which have arisen from the advent of mathematical logic and the digital computer, leading to the idea of sound computation.

However it is the forward looking aspects which are the more important. This will include discussion of the step forward from “sound computation” to “crisp AI”, making use of the idea of “FAn oracle” a refinement of the analytic oracle.

In the architecture which we later propose, proof and computation are intimately intertwined. Throughout their history these two concepts have frequently been transformed, and have frequently impacted upon each other. Far from stabilizing it seems that the pace of change in our ideas about how these concepts interrelate has accelerated.

The chapter begins historically, but ends in the contemporary context with some technical points intended to open up new possibilities for the relationship.

The history of computation goes back a long way. Our history begins with the idea of proof in mathematics, which we first know of in the beginnings of Greek philosophy and mathematics with the philosopher Thales. Whether the reasoning at this point could properly be called deductive is moot. Two later landmarks which may be thought of as pinnacles in the articulation and application of deductive methods were the logical writings of Aristotle’s Organon and the mathematical tour de force of Euclid’s Elements. These works became the standard for two millennia in philosophical logic and in deductive rigour and axiomatic method in mathematics. Substantial advances beyond these would not be made until the 19th century.

5.1 Before the Greeks

It is generally said that the ancient Greeks invented mathematics and deduction.

This is probably true of the self conscious use of deduction and the development of mathematics as a theoretical discipline.

However, there are some points worth making about what happened before that. The man in the street, hearing the word “mathematics” thinks primarily of arithmetic, and of practical skills of computation. He may have little conception of mathematics as a theoretical discipline.

Mathematics in that sense preceded the Greeks. Notation for numbers, numerical computation and elementary geometrical capabilities probably go back as far as 10,000 years BC. Before classical Greece, the principal sources of expertise in these practical matters were Babylonia and Egypt. There are at this stage methods for computation and for solving certain kinds of algebraic problems, but these methods are presented without any attempt at justification, there is neither a theoretical discipline establishing results by deductive proof.

The ability to perform elementary steps of deductive reasoning is probably coeval with descriptive language, since one cannot be said to understand the language without grasping and being able to apply elementary conceptual inclusions. That is to say, for example, that if someone knows the meaning of “azure” then he knows that everything which is azure is also blue. However, the first signs of deduction being used elaborately and systematically, and the idea or deductive proof are
found at the beginning of the classical period of ancient
Greece at about 600BC.

The Greek civilization is thought to date back to
about 2,800 BC, but mathematics as a theoretical dis-
cipline originates with the school of the philosopher
Thales in Ionia at around 600 BC.

5.2 Sketch

The idea here is to sketch the history of proof, and the
history of computation.

Initially these two seem quite separate.

The potential for connection begins with the ap-
proach to formality found in Aristotle’s Aristotelian
syllogistic, and with the axiomatic method for mathemat-
ics documented in Euclid’s Elements. Aristotle’s work
dominated the philosophical study of logic until the
19th century, but his logic was the object of philosophical
enquiry rather than a practical tool in deduction.

The axiomatic method on the other hand had proved
highly effective in Greek geometry. The standards then
set were held as an ideal in mathematics, but were not
matched in subsequent developments in mathematics,
even (or especially) when mathematical innovation and
the effectiveness of mathematics as a tool for science and
engineering was scaling new heights, until we come
to the end of the 19th century.

Leibniz connects Aristotelian logic, by conceiving
a scheme for arithmetisation of Aristotelian logic,
which he thought would permit the truth of a subject/predicate proposition when arithmetically coded
in his universal characteristic. Calculating machines
were already in existence, and Leibniz designed such a
machine contributing to the development of the tech-
nology. Important though these ideas were, they were
neither applicable nor influential for hundreds of years.

The next important developments both occur in the
19th century. The first was the design by Babbage’s analytical engine, the first conception of a universal computational engine. The second
was the idea of formal proof first exhibited in Frege’s
Begriffsschrift [Fre67]. Frege’s idea was the first advance in rigour beyond Euclid’s axiomatic method, adv-
cancing that method by requiring not only that that the
axioms, postulates and definitions on which a proof is
based are made explicit, but also by requiring that the
rules by which conclusions are drawn at each step in a
proof from these premises and previous conclusions, be
carefully specified in advance. This required that the
language in which the proofs are conducted be formally
specified.

There is a connection here, not explicitly made, be-
tween these two developments. By designing a uni-
eral analytic engine Babbage had implicitly defined a
notion of computability. A function is computable if it
can be calculated by Babbage’s analytical engine. By
requiring that the rules of inference used in a proof be
fully specified Frege was demanding that the correct-
ness of such a formal proof be mechanically checkable,
that in principle the correctness of such a proof could
be verified by Babbage’s analytical engine.

Though we can now see this connection, it was not
made explicit for a while.

Subsequent to these two developments the notion of
proof and of the axiomatic method were refined as a
part of an explosion of work in logic and its application
to the derivation of mathematics. A key figure in
this was David Hilbert, who began by advancing the
axiomatic method, for the first time achieving levels of
rigour surpassing those of Euclid, following the initial
insights into the nature of proof by Frege. He did this
by returning to the axiomatic theory of geometry and in
the process found and resolved weaknesses in Euclidean
geometry. By this time Frege’s logical system in
which Frege had sought to formally derive arithmetic
using his new logical methods had been found to be in-
consistent (by Bertrand Russell), precipitating a sense
of crisis in the foundations of mathematics. Hilbert
responded to this with a program for putting mathe-
ematics on a firm foundation, and a major part of this
was a new domain of research which he called meta-
mathematics.

This new domain of meta-mathematics was the study
of the kinds of formal system in which mathematics
might be developed, with a view to establishing the
consistency of such systems.

Bertrand Russell came up with a new logical system
which he called the “theory of types” for use in his
collaboration with A.N. Whitehead on Principia Math-
ematica. With this formal logical foundation they car-
ried through the formal development of a significant
part of mathematics, starting (after more logical and
abstract considerations) with arithmetic. It is notori-
ous that in that monumental work they prove the ele-
mentary arithmetic theorem that 1 + 1 = 2, do not reach
that point until page 362. This emphasises an impor-
tant difference between computation and proof as it
was then understood. It is possible to prove theorems
which express the results of arithmetic computations, but the effort and complexity involved in such proofs outstrips by many orders of magnitude the difficulty of simply performing the calculation, which would be accepted in an informal proof which depended upon it without further justification. Formal proofs are so detailed that their complexity compares badly with the informal proofs which were then and still are the norm in mathematics, and in the case of elementary equations with the difficulty in performing arithmetic by the usual methods.

5.3 The Status of Proof

Early in the 20th century radical changes have taken place in the conception of proof and of the role of proof in mathematics.

There were at this time several distinct attitudes towards proof in mathematics which are of interest here.

The simplest is that a formal deductive system consists of a collection of axioms which are simple enough to be self-evidently true, together with some rules of inference which likewise may be thought self-evidently to preserve truth. In consequence all theorems derivable from the axioms using the rules of inference will be true.

The discovery of the inconsistency of Frege's Grundgesetze undermined this position by showing that a system whose soundness in the above sense seemed self-evident, might nevertheless prove inconsistent.

In devising a new logical system intended to avoid the kinds of difficulty found in Frege's, Russell was therefore not able wholeheartedly to claim soundness on these grounds, and instead suggested a partly post-hoc rationale. Russell thought that the success of a logical system in deriving mathematics without proving any contradictions provides evidence that the system is sound.

The developments in mathematics which rendered mathematical analysis rigorous by eliminating the use of infinitesimals lead on to a liberal conception of the notion of function as an arbitrary graph which proved unacceptable to some mathematicians, notably Leopold Kronecker. Kronecker's scepticism about the new mathematics particularly as it is found in Cantor's Set Theory, was an early example of constructivism in mathematics in which mathematical objects are required to be constructible by limited means resulting in an ontology radically less expansive than that of set theory. Associated with this more frugal ontology there is a completely different attitude to proof. The effect is not just that mathematical proofs are expected to use more limited means, but that the subject matters of mathematics are modified because of ontological reservations, and in effect certain domains of enquiry are banished (of which again, set theory as conceived by Cantor is a principal exemplar).

Against this "impoveryment" of mathematics the mathematician David Hilbert reacted. He devised a program intended to establish the soundness of the whole of classical mathematics to a standard equivalent to that of constructive mathematics, and introduced the concept of meta-mathematics to that end. Meta-mathematics is the study of the kinds of formal deductive system which can be used in the derivation of mathematics. Hilbert's aim was by developing metamathematical techniques (now known as proof theory) to establish the consistency of the whole of mathematics by means which would be acceptable to constructivists.

We have now three attitudes toward proof.

The first consists in choosing a deductive system carefully, adopting some underlying philosophical rationale (in Russell's case this was the avoidance of "vicious circles"), choosing axioms and rules guided by the chosen rationale, and then seeing whether the system proves to be adequate for the mathematics and seems to be immune from paradox.

The second rejects the idea of a post hoc justification, is ontologically conservative, and expects a close connection between the demand for a constructible ontology and the kind of proofs which are acceptable.

The third seeks a bridge between the two, with the aim of allowing mathematics the scope which it might assume under the first approach, while achieving the standards of constructive rigour sought in the second approach. This promising idea turned out to be unrealizable, but its failure spawned another, in which relative consistency proofs provide a partial ordering of formal deductive systems providing a measure of strength and of risk.

5.4 Incompleteness and Recursion Theory

The next important connection between proof and computation comes from Gödel in connection with Hilbert's program, and is the first of a series of advances relevant
to proof and computation which appeared in the 1930s and 40s. Gödel’s famous incompleteness theorems were obtained by the technique of arithmetisation of which we saw a precursor in Leibniz’s *calculus ratiocinator*. The result demonstrates a limitation on what can be done with certain kinds of consistent formal deductive system, viz that no such system can prove every statement of arithmetic. The precise definition of the kind of system within the scope of the proof provides formally the connection implicit in the idea of Frege that the notion of proof be *computable*.

We are now on the verge of an explicit theory of computability. The idea of computability first appearing implicitly in the design of Babbage’s analytical engine, now comes under theoretical scrutiny in the new discipline of mathematical logic. Several different formal conceptions of *effective process* are enunciated (by Church, Kleene, Post, and Turing) and are shown to be equivalent in expressiveness, leading Alonzo Church to put forward *Church’s thesis* that these equivalent formal ideas capture the informal notion of effective calculability.

### 5.5 Computing Machinery and Proof

In the meantime, the technology of computation has moved on. The analytic engine designed by Babbage was a mechanical computer, and its realization was beyond his resources. Computation was by the 1940s in transition from electro-mechanical to electronic technology, universal computers were now within reach.

When they arrived, though predominantly applied to the purposes of businesses or for scientific calculations, they were soon also applied in academia for the construction and checking of formal proofs. The new academic discipline of Computer Science and the rapidly growing computer industry became principal users and developers of formal languages, in which the algorithms to be executed by computers were described. The theory behind computing was also studied, and the techniques of mathematical logic extensively adopted. Mathematical methods were advocated for proving correctness of programs, and even of the “hardware” which executed the programs, and the expected complexity of these proofs encouraged the development of software to assist in constructing and checking the proofs.

These developments spawned new kinds of logical languages, new approaches to reasoning in those languages, and new applications for proof. Some of these are significant here because the notion of proof for which our project is to support and exploit is distant from those which predate the computer and computer science.

A principal influence in shaping a more diverse conception of proof is the development of automatic and semi-automatic or interactive theorem provers. Three different kinds of such software systems are of interest.

The first kind is the one closest to the new conception of formal proof which came from Frege. In such a system a proof is a sequence of theorems each of which is either an axiom of the logical system or is derived in a specified way according to an inference rule of the logic from one or more theorems appearing earlier in the list. The end result of the proof is the last sentence in this sequence.

A second approach to theorem proving simply leaves the details to the software, which is not required to construct or verify a detailed formal proof, but instead is written to check reliably for derivability in the relevant formal system and to deliver an informal proof of the result.¹

The third approach which is of greatest relevance here began again at Edinburgh University as a theorem prover for a Logic for Computable Functions (LCF). This approach to proof later became known as the LCF paradigm. It makes use of a feature of certain kinds of programming language which is called an “abstract data-type”, whereby a new kind of structure is defined which can only be constructed by the use of a limited set of functions defined for that purpose. The implementation of such a theorem prover begins with the specification of a formal deductive system, and proceeds by the implementation of an abstract data type which corresponds closely to the deductive system. Each of the available constructors in for the abstract data type is required to correctly implement one of the axioms or inference rules of the logic delivering only theorems which could be obtained using the relevant axiom or inference rule.

By this means it is ensured that values which can be computed using this abstract data type must be theorems of the logic, for the structure of the computation

---

¹The classic example of such a prover is NQTHM, a theorem prover first developed at the University of Edinburgh by Robert Boyer and Jay Moore, others in this genre include PVS from the Stanford Research Institute, and Eves, developed by a small Canadian company.
yielding the value corresponds precisely to the structure of a proof in the logic.

In this approach to the automation of proof, the proof itself as an object has disappeared, instead of a sequence of theorems of which the result demonstrated is the last, we have a computation which by means guaranteed sound delivers as its result the desired theorem. Here the proof is a computation.

The implementation of the abstract data type provides a logical kernel, which ensures that any value which can be computed of type \texttt{thm} will be a theorem of the logic. This is by itself too primitive to be useful for a user of the theorem prover. More complex software is developed which partly automates the computation of the theorem (which is analogous to and serves as the discovery of a proof). These additional layers of software are then made available to the user through an interactive programming interface through which he can undertake proofs, possibly augmenting the proof capability by further programming of proof discovery methods.

Several further steps along this direction are significant for us here. As with the kind of formal proof found in \textit{Principia Mathematica} complexity remains an issue. Even with a computer to do the work, the search for a formal proof, or even the checking of the proof can be onerous. For the sake of efficiency it is therefore usual to implement in the logical kernel of a theorem prover following the LCF paradigm a number of derived rules, the use of which significantly improves the computational efficiency of the theorem prover without impairing its soundness.

Taking this a step further we may consider the addition of any inference rule which is known only to perform inferences which could have been done in the primitive logic. Such enhancements in principle improve computational efficiency without alteration to the theorems which are provable in the system.

### 5.6 Sound Computation as Proof

The difference between computation and proof may be seen as a difference in perception or interpretation.

A computation is an operation on data yielding data, an inference is an operation on propositions yielding a proposition. Propositions are \textit{interpreted} data. The data transformed by a computation may have multiple possible propositional interpretations, relative to which the computation might be seen to be a sound logical inference.

A computable function becomes a sound inference rule if can be given a suitable specification. This is always possible, though this nominal possibility does not necessarily translate into something useful.

### 5.7 Oracles and the Terminator

We know from recursion theory that there are unsolvable problems. One of these is the Turing machine halting problem, the problem of determining for any given specification of a Turing machine and its tape whether the Turing machine will halt.

Recursion theory studies degrees of unsolvability using the notion of an Oracle. An Oracle is a hypothetical machine which answers a problem which is not in fact effectively decidable, such as the Turing machine halting problem. This idea can be used to give an upper bound on what one could conceivably achieve by way of problem solving capability in a program running on a digital computer.

The halting problem is equivalent to the problem of deciding whether a sentence is a theorem in the kind of deductive system which is suitable as a formal foundation for mathematics. Automatic proof methods for first order logic can be arranged to deliver with the proof of an existential theorem a term which is a witness for that theorem, which satisfies the body of the existential. If a design problem is expressed as existential proposition in which the body stipulates the conditions required of a solution to the design problem and a solution to the problem is any term which satisfies the body of the existential, then degree of recursive solvability corresponds to a design capability.

### 5.8 Self Modifying Procedures

Because of their intelligence, people can solve problems which have no feasible algorithmic solution. It doubtful that they do this because intelligence can draw upon resources which go beyond the limits of Turing computability, that possibility and its consequences are outside the scope of this book.

Let us assume that, as Turing himself believed, there is nothing in human intelligence which might not be realized in a Turing machine. That entails, that for any particular definite problem domain, there exists an algorithm which performs as well in that domain as an
above average appropriately trained and experienced human being.

The algorithm might not be one which any of us will ever be able to code. It is possible that its complexity, like that of the human brain, might exceed anything which a human beings could design and implement. It might be that such an algorithm could only arise through a process of evolution.

Our project is intended to support the use of evolutionary processes to realize machine intelligence in solving crisp problems. Some details of how that might be realized are touched upon here.

Let us suppose that some implementation or extension of our architecture makes provision for its own evolution over time. Because the notion of proof is assurance sensitive, some assurance may be sought that such changes do not render the system unsound and thus degrade the assurance which we could attach to its results.

It is possible to implement a system which allows for self modification in a qualified way, permitting only modifications which preserve the trustworthiness of the system. This can be done by admitting only modifications which preserve that level of assurance, which they will do if they are conservative.

For any formally precise property of systems, there will be a system which undertakes self modification in ways which are guaranteed to preserve that property.

The implementation of such a system is not difficult, for any proof tool which implements a reflection principle of the kind discussed above will be a self-modifying proof system which preserves the property of infallibility in asserting theorem-hood. No simplistic approach to realizing intelligence in such a way would be likely to succeed.
Chapter 6

Rigour, Scepticism and Positivism

[The main point of this chapter is to lead us into two features of positive philosophy and of the architecture for crisp AI. These are:

Epistemic Retreat At its simplest this is the idea that, instead of asserting a claim, one retreats to describing the evidence in favor of it. For science it is the idea that a scientific theory is to be presented as an abstract model, and rather than asserting the “truth” of the model (whatever that might mean) one makes statements about its fidelity and utility in various circumstances. Part of this is the idea that the assertion of a logical truth is to be done formally, whereas empirical claims are asserted only informally.

Graduated Scepticism Not sure whether to call this scepticism, but the idea is that instead of choosing between two theories and then calling one true and the other false, one suspends judgement about truth and confines oneself to comparative evaluation (of various kinds).

Assurance and Authority The second principle idea, which is applicable primarily to logical truth, is that these truths are asserted by authorities in which we may think of an authority as anything which may wish to formally express an opinion about the truth of some conjecture. Authorities may do so baldly, or may come to such an opinion in the light of the opinions of other authorities. In the latter case, the assertion will mention those other authorities, and the combination of the set of authorities thus cited and the authority expressing the opinion is regarded as a level of assurance which fits into a lattice structure providing a partial ordering of such assurance levels.

The chapter also needs to connect to contemporary reasons for scepticism and to the reasons for doubt about philosophy arising from the work of Quine and Tarski.

] In defending key elements of the philosophy with which I propose to underpin the automation of reason I have found it necessary to reject wholesale two of the most potent influences on analytic philosophy since the mid 20th century.

If philosophy can go so badly astray, even after the modern revolution in logic transformed it, what hope is there that the ideas I champion can be better. Is the descent into nihilism not inevitable?

This line of thought is not new, it recurs throughout the history of Western Philosophy. In this chapter I trace some of this history. I do this partly to make my negative conclusions seem less startling and more plausible, but principally because key features of Positive Philosophy (the positivism!) have evolved from more extreme scepticisms over a long period of time, and are best understood in that context.

In each historical era different sources of dogma provide new principal targets for the sceptic. In the first wholly sceptical philosophies, those of the Pyrrhonists and the sceptics of Plato’s academy, a principal target was the dogmatic metaphysics of the pre-Socratic philosophers, and the great syntheses of Plato and Aristotle. When pyrrhonism first reappeared in Europe its target was the dogmas of Catholicism. Today we have many sources of dogmatic disinformation, but the one of greatest concern here is academia in general and analytic philosophy in particular.

The thought of the radical Greek sceptics was passed into history principally in the comprehensive sceptical writings of Sextus Empiricus[Emp33], which began to
influence modern European thought after being translated from Greek into Latin in the sixteenth century.

These revived sceptical arguments figured at first in the religious controversies of the reformation, and were deployed against Catholic dogma. Influential among these new pyrrhoneans was Michel de Montaigne.

Once these sceptical arguments were rediscovered they were not so much adopted as moderated. This was particularly important for science, and it was scientific philosophers such as Gassendi and Mersennes who sought a constructive scepticism compatible with the new science. The most important enduring tradition which emerged from this moderation was positivism, which rejects pseudo-science and promotes higher standards for “positive” science.

In modern times the arch sceptic and the target of much anti-sceptical argument in epistemology has been David Hume. Greek scepticism, despite providing most of the ammunition for Hume, figures less prominently. Hume’s philosophy was a sequel to a variety of sceptical thought through the sixteenth and seventeenth centuries, which begins with the religious controversies of the reformation.

6.1 Systematic Skepticism

Systematic, wholesale skepticism, the doubt that there is any true knowledge, was found in two important philosophical traditions in post-classical Greece. The best known of these was that associated with the name of Pyrrho of Ellis (though it is not known how many of their ideas actually came from Pyrrho). The other is the Academy of Plato, which, after the death of Plato underwent periods of radical scepticism.

The pyrrhoneans offered a comprehensive range of skeptical arguments which were later documented by Sextus Empiricus, which we need not examine in detail. We need only consider a single kind of very general argument, which is argument by regress of justification. In its simplest form it asserts that:

1. to count as knowledge a belief must be justified

2. a justification of a proposition consists of a number of known premises, together with an argument known to be conclusive which shows that the conclusion must be true if the premises are.

3. a justification therefore depends on prior knowledge, either of the supposed evidential support or of the validity of some form of inference, and we therefore have an infinite regress and no knowledge is possible.

Of course, this argument undermines itself, and so cannot itself yield the knowledge that knowledge is impossible.

Philosophies which are skeptical about knowledge risk incoherence, and are often accused of inconsistency. The most naive incoherence would be to claim to know that nothing can be known. This accusation has been levelled, but it is doubtful that any skeptic on record made that mistake. The closest is the doctrine that we can know nothing but this single proposition, which is consistent if a little ad hoc.

There are two subtler kinds of inconsistency which are more significant for us.

The first is equivocation about the meaning of the verb “to know”. We are concerned with “knowing that”, knowledge of true propositions, rather than “knowing how”. This generally involves a true belief, and usually depends on the knower having adequate grounds for his belief, thus the formula: knowledge is justified true belief. However, the requirements in terms of justification vary widely. In some discourse this requirement of justification lapses, and someone may be said to “know” a fact simply because he has been apprised of it. On the other extreme the requirement for justification can be exhaustive, in some contexts the required justification must be conclusive. What counts as conclusive is also up for grabs. In the context of a philosophical discussion a justification may need to be the possession of evidence which logically entails the candidate proposition.

Arguments in favor of radical scepticism will often implicitly assume the need for the highest standards of justification. However, once established in this way, the resulting scepticism may be applied to all kinds of “knowledge” even in contexts where the required standards of justification is weak or nugatory.

One way of preventing inconsistency is to oppose dogmatism with doubt. Thus, the skeptic responds to a claim to knowledge, not by contradicting that claim, but merely casting doubt upon it. This leads to sceptics who seek to establish “equipollence” a situation in which the evidence for and against a proposition are perfectly balanced.

Pyrrhonean scepticism advocates reserving judgement, keeping an open mind, rather than the dogmatic assertion of some proposition. The pyrrhonean is sometimes portrayed as someone seeking knowledge who fails
to find the certain knowledge that he seeks. There are two elements of inconsistency or equivocation which can be observed in an account at this level. The first is the contrast between the two objectives, on the one hand that of seeking knowledge (albeit unsuccessfully), and on the other of seeking equipollence in respect of any particular judgement. Surely if equipollence is an objective, then this represents a prejudice against the possibility of knowledge, rather than a positive attitude of seeking knowledge? This accusation can perhaps be deflected by presenting the search for equipollence as a way of testing a conjecture, and hence establishing the proposition by failing to show equipollence.

A second point of apparent equivocation is in the word dogmatic. In ordinary parlance a dogmatist is someone who holds onto a fixed belief in the face of contrary evidence. In Pyrrhonean usage one becomes a dogmatist in virtue of even the smallest departure from complete doubt. The mere expression of a tentative belief counts as dogmatism. The sceptical arguments however, very often depend on the demand for conclusive justification, and hence have force only against dogmatism in its more ordinary usage, rather than in the more liberal interpretation which embraces mere opinion.

When we come to positivism we find a kind of mitigated scepticism in which the systematic doubt is a part of scientific method, is a way of testing hypotheses. Elements of skepticism provide a basis for positivism and its notion of positive science, and will have an influence on our proposed “architecture of knowledge”, so we seek here a way of removing inconsistencies and equivocations to obtain a defensible position on which to build.

In this connection we leave the Pyrrhonists to consider the graduated scepticism of the Academic skeptic Carneades.

### 6.2 positivism

- Positivism as a moderation of skepticism
- Positivism as scientific method
- Kolakowski’s characterization of positivism
- Hume’s fork

The term positivism was coined by Auguste Comte and refers to the whole of his broad ranging philosophical thought. Typically however the term has been used for a narrower position centering around his conception of positive science, and has been construed as relating to a philosophical tradition many elements of which were first clear in the philosophy of David Hume. This is the perspective on positivism which is found in the history by Kolakowski[?], and which we work with here.

Comte held that the human mind progresses through distinct phases, the theological the metaphysical and the positive. In the theological stage the hidden nature of things is connected with the belief in supernatural beings, in the metaphysical stage the hidden nature of things is sought without resort to the supernatural, and in the positive stage scientific understanding supersedes both theology and metaphysics through the formulation of universal laws about the phenomena which do not depend on hidden entities. He offered his conception of positive science not as an innovation, but rather as an account of the scientific method already established by his predecessors, men such as Bacon and Galileo.

### 6.3 Rudolf Carnap

#### 6.3.1 Tolerance, Pluralism, Metaphysics

These three topics are intimately interwoven in Carnap’s philosophy. I’m going to attempt an elucidation of elements of Carnap’s philosophy by discussing some interpretations of Carnap on pluralism which seem to me to be, in various degrees, mistaken.

I discuss the views on Carnap’s pluralism of three fictitious philosophers, R, A and K. Their positions are suggested to me by the writings of three actual philosophers, but it is not necessary for my present purposes to resolve the question of what those real philosophers actually meant, it suffices to discuss positions they might possibly have meant.

First a brief preliminary statement of some key aspects of Carnap’s notion of pluralism and his principle of tolerance. According to Carnap’s intellectual autobiography [Car63a, Car63b] his tolerance was anticipated in his student days by a willingness to discuss issues with friends in a variety of “philosophical languages”, which corresponded to distinct and incompatible metaphysical stances. The principle of tolerance did not appear until much later, in “logical syntax”
It is understandable that people will come away from this book with diverse and incompatible ideas of what Carnap’s pluralism is. Carnap does not actually use the word pluralism in the book. He does enunciate his “principle of tolerance” in

6.4 1

7:

It is not our business to set up prohibitions, but to arrive at conventions.

Which is further explained in that section as:

Everyone is at liberty to build up his own logic, i.e. his own form of language, as he wishes.

I begin with R. R claimed that he was more pluralistic than Carnap. I reacted against that, since I didn’t see that Carnap’s pluralism actually excluded anything. The feature of his pluralism which he thought beyond Carnap’s was that he would accept that the same sentence could have different meanings and different truth values. But this is also the case for Carnap. For Carnap this would be possible by having that same sentence in two different “language frameworks” (which we may perhaps think of as being both the grammar of the language and the semantics in some form). I mention this here because it contrasts with the next one.

K took a more substantial interest in Carnap’s philosophy and wrote quite a bit about it. He recognizes two interpretations of Carnap’s pluralism, on the one hand as a substantial thesis, and on the other as a proposal. However, his critique is concerned with the former. In this case he takes Carnap to be taking a radical view entailing the indeterminacy of truth in languages like arithmetic and set theory, so that Carnap is to be understood as saying that these truths are not completely determinate but can be chosen arbitrarily.

<table>
<thead>
<tr>
<th>Year</th>
<th>Figure</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>585 BC</td>
<td>Thales</td>
<td>deductive mathematics</td>
</tr>
<tr>
<td>6th C BC</td>
<td>Ionia</td>
<td>metaphysical cosmology</td>
</tr>
<tr>
<td></td>
<td>Zeno</td>
<td>paradoxes of motion</td>
</tr>
<tr>
<td>5th C BC</td>
<td>sophists</td>
<td>man is the measure of all things</td>
</tr>
<tr>
<td>470-399 BC</td>
<td>Socrates</td>
<td>the organon and the metaphysics</td>
</tr>
<tr>
<td>384-322 BC</td>
<td>Aristotle</td>
<td>the theory of forms</td>
</tr>
<tr>
<td>428-348 BC</td>
<td>Plato</td>
<td>the axiomatic method</td>
</tr>
<tr>
<td>c300 BC</td>
<td>Euclid</td>
<td></td>
</tr>
<tr>
<td>c365-270 BC</td>
<td>Pyrrho</td>
<td>scepticism</td>
</tr>
<tr>
<td>c180-110 BC</td>
<td>Carneades</td>
<td>graduated scepticism</td>
</tr>
<tr>
<td>c1288-1348</td>
<td>William</td>
<td>nominalism</td>
</tr>
</tbody>
</table>

Table 6.1: Rigour, Scepticism, Positivism
Chapter 7

Epistemic Retreat

This chapter is philosophically epistemological, providing a kind of constructive epistemology in the form of some architectural principles in relation to the representation of knowledge in networked storage systems and its processing by distributed processing systems involving both natural and artificial processing elements.

The chapter is therefore a conflation of aspects of philosophical epistemology with abstract architectural design for knowledge processing systems. This necessitates (or flows from) some novelty in my conception of both these enterprises, so I will begin with some discussion of the innovations involved.

Epistemology may most briefly be characterized as the theory of knowledge. It has typically been anthropomorphic, i.e. concerned specifically with human knowers, and sometimes linguistic, concerned with the specifics of the meaning of the concept of knowledge. Here I am interested in knowledge in information systems, not in human brains, and seek to avoid giving any deep scrutiny to the meaning of the word know. The conduct of epistemology without attachment to the concept of knowledge is analogous to the preference in science for avoiding vague and relative concepts such as “hot” and “cold”, in favor of objective and precise properties such as “temperature in degrees Kelvin”. Instead of claiming knowledge, we will be aiming to provide more objective descriptions of grounds for the truth of propositions, or of evidence showing the reliability and fidelity of abstract models of the real world. I do not prescribe particular measures, but provide instead a context in which a plurality of comparative evaluations can be accommodated.

The first comparison is the one provided by Hume’s fork, and this has a major impact on the epistemology and the knowledge architecture. Because of the great precision with which “relations between ideas” can be expressed, such propositions we regard as assertable, and in connection with such assertions “epistemic retreat” involves a form of assertion in which the grounds for asserting truth are made explicit. On the other side of Hume’s fork, in relation to “matters of fact”, epistemic retreat in the first instance reflects the imprecision in our knowledge of Plato’s fleeting world of appearances. Scientific knowledge is regarded as embodied in abstract models of the real world, which are applied by deduction (this is our version of a nomological-deductive scientific method).

A principle effect of the status of Hume’s fork in this epistemology is that our formal knowledge base asserts only logical or analytic truths. In it, scientific theories are represented through abstract models, which will in general have concrete interpretations, but which have the logical status of definitions rather than assertions. Epistemology has been concerned with the refutation of scepticism. It here embraces an open scepticism (but not a dogmatic negative scepticism), and is therefore oriented not with the refutation of scepticism but with the establishment of a viable constructively sceptical system.

The first step in the moderation of a pyrrhonian scepticism is the recognition that, though no proposition is absolutely certainly true, some appear to be more certain than others, and among the more doubtful there are also degrees of doubt. Can we know with absolute certainty that some proposition is more certain than another? No, but we can form opinions about relative certainty which are themselves more solid than our opinions about truth (again, abstaining from absolutes). Among these comparative assurance judgements are the relations between a proposition and the evidence we have for it. It will generally be the case that our confidence in the evidence on which we base an opinion
will by stronger than that in the proposition we infer from the evidence. Similarly, it will be the case that a statement effectively describing the evidential support for a theory will can be more confidently asserted than the theory itself.

This is the principal way in which Metaphysical Positivism interprets the positivistic principle that science should not go beyond presentation of observational data. It is not taken to impede the formulation of general scientific theories which go beyond the content of any possible body of experimental evidence. It is taken instead to impede the assertion of such empirical generalizations as truths. The positive scientist instead compiles various bodies of evidence which provide a basis for decisions about when the scientific generalization might be applied.

The experimental data obtained will provide information about the fidelity and accuracy of the theory as modelling the real world in a variety of circumstances, so that someone contemplating use of the theory will be in a position to form an opinion about whether the theory will provide a sufficiently reliable and accurate model for his purposes.

The mitigated scepticism of positivistic philosophy, following David Hume, accepts a priori truths of reason as certain, but expects of positive science that it does not go beyond what is entailed by the experimental or observational evidence. Since most scientific theories involve empirical generalizations which go beyond any finite collection of particular observations, positive science would seem by this doctrine to be eviscerated.

Metaphysical positivism recognizes two principles which constitute a kind of foundationalism. The first is connected with the sceptical doctrine that all we know is that “appearances appear”. However, what we count as an “appearance” is not confined to sensory impressions. Any impression which we may form on the truth of some proposition, whatever its source, is counted as an appearance. Such appearances we accept as what they are, and the body of science is considered to constituted just an organized presentation of a great deal of such material. The enterprise is not solipsistic, it is collaborative, and we therefore recognize as significant the source of the impression, the identity of the person or entity who was the subject of the impression. Because of the broadening of the notion of appearance, propositions expressing such appearances are called “opinions” and are to be tagged or digitally signed by the entity whose opinion they are. The kind of entity which has opinions is called an “authority”.

Because of the collaborative nature of the enterprise, an opinion will normally be formed by an authority on the basis of (as entailed or otherwise supported by) some collection of opinions of other authorities. It is therefore normal for an opinion to be expressed, on the assumption that various authorities can be trusted, or more specifically, on the assumption that all their previous opinions are true. The collection of authorities on the basis of whose opinions a further opinion is formed, together with the authority forms the new opinion, give a measure of the risk associated with the opinion which which I call a degree of assurance. These degrees of assurance are partially ordered. The more authorities whose opinions are involved, the lower the level of trust. The partial ordering becomes a lattice when we allow that an opinion may be endorsed by several authorities on the basis of distinct collections of other opinions. Adding more independent opinions increases the degree of assurance.
Chapter 8

Language Planning

In his intellectual autobiography [Car63a] Rudolf Carnap provides, as well as a sketch of the development of his thinking, a section covering ten important topic headings under which his ideas, aspirations, and accomplishments in these areas were discussed. One of these topic areas is “language planning”.

Under this heading Carnap mentions two different but related problems, the problem of constructing “an auxiliary language for international communication”, and that of constructing “language systems in symbolic logic”. In both cases he credits Leibniz’s previous contributions.

Perhaps because different aspects of the latter problem were a major part of his work and were addressed under different headings (such as logical syntax and semantics) he devoted most of the space under this heading to non-symbolic languages.

He mentions the shift in his perspective from the universalist conceptions of logic found in Frege and Russell to his own pluralism, and his concern with the difficulties arising in the choice of languages for science and the systematization of science in a pluralistic context.

Though Carnap devoted much time to devising and describing methods for defining languages, their semantics and deductive systems, the problems which might arise from the use in science of a plurality of such formal languages were not addressed.

Since his day as a result of the invention of the digital computer there has been an explosion in the use of formal languages, and other kinds of representation of data or knowledge.

If we wish to reason using the information held in such representations, then we embrace a pluralism broader than Carnap could have conceived, and something like Carnap’s problem of “language planning” in its formal aspect, becomes pressing.

This problem is considered here as a prelude to the discussion of architectures for formal analysis in which this diversity of representation can be encompassed without compromise to logical coherence.

8.1 Languages, Notations and Representations

The support of rigorous deductive reasoning is of the essence both in the projects of Leibniz and Carnap and in the successor we are considering.

Leibniz’s conception of how propositions should be expressed in order for his calculus to operate was radical even by today’s standards. The form of proposition was the subject predicate form, of which there were four variants, and in which the subject and predicate were each represented arithmetically as a pair of co-prime natural numbers. His calculus involved quite simple arithmetic operations on these numbers to determine whether the proposition was true or false.

The merits and limitations of this proposal are not our present concern, it is mentioned to emphasize the very broad limits on the kinds of representation and proof which are open to consideration.

By contrast with Leibniz, our aim is to admit into our architecture a very broad range of representations rather than devising one representation suitable for the calculus and expecting all knowledge to be coded using that method of representation so that the calculus could be applied.

The general plan is that any representation be admissible provided only that it has a definite significance, preferably a formally specified significance or semantics.
8.2 Universalism and Pluralism

Frege and Russell, pioneers of the formalization of mathematics, each conceived of a single language in which the whole of mathematics might be expressed and proven.

Frege’s formula has been paraphrased:

Mathematics = Logic + Definitions

Which may be read as asserting that given a suitable logic, the whole of mathematics can be developed simply by writing down formal definitions of the concepts of mathematics, and then proving the required mathematical theorems involving those concepts.

There is no reason in this scheme even to contemplate the possibility of making use of more than one language, but if we do, problems immediately arise. The simplest is that the idea of a formal deductive system is such that deductions take place exclusively in one language. The deduction rules which govern the structure of proofs in such a system only allow premises within a single language to be used in a proof. The effect of this is to ensure that if more than one language is used, the results obtained in one language cannot be used in the proof of results in another language. This kind of awkwardness is not a feature of informal mathematics and would represent a serious and avoidable flaw in a formal approach to mathematics.

Rudolf Carnap, inspired by Bertrand Russell to adopt a “scientific” approach to philosophy, and hoping to do for science what Russell had done for mathematics, i.e. to establish methods suitable for the formalization of science. This was to be done by adapting the new logical methods to the domain of empirical science. There was in this an immediate difficulty, arising from the necessity of making empirical claims. Concepts applicable to the material world cannot in general be defined in terms of purely logical contexts. It does not seem that one can proceed by analogy with Frege and plausibly argue that:

Empirical Science = Logic + Definitions

Something more seems to be needed, and it is natural to suppose that this is a language which goes beyond logic by including concepts which are empirically significant.

So it appears that Carnap’s enterprise compelled him to consider new languages. Even before this, Carnap already was, by his own autobiographical account, a pluralist with respect to language (though he might not at that time have described himself in those words). His pluralism at that time was primarily in relation to distinct ways of talking about the world which were associated with different metaphysical postures. The most solid example of this is the distinction between the materialistic language of science, in which material objects are the subject matter of the language, and the phenomenalistic language of empiricist philosophers going back to Hume, who confine themselves to talking not about material objects but about the sensory evidence which we have of that world.

It is characteristic of Carnap’s quite original attitude to metaphysics that he was happy to talk in either of these languages even though their proponents took them to be representative of incompatible views on metaphysics. Carnap regarded the underlying metaphysical issues as meaningless, and took a pragmatic attitude to the use of the languages, eschewing any supposed metaphysical commitment that might be supposed to entail. His pluralism, to be made explicit later in his “principle of tolerance”, was to the effect that any language can be adopted, subject to pragmatic considerations, rejecting the metaphysical dogmatists who asserted the legitimacy of only that language which embodied their preferred metaphysics.

At the same time as Carnap enrolls himself into what he thinks of as Russell’s program, in the 1920’s, the high noon of the Vienna circle, Hilbert at the center of a group of mathematicians in Berlin is moving forward on a different conception of how to formalize mathematics, and Carnap, eager to absorb all the new developments in logic which might be relevant to his project, is paying attention.

Hilbert’s approach to formal mathematics, though fully embracing the new logical methods and pressing them forward, is distinct from the universalistic conceptions of Frege and Russell and harks back to the axiomatic method first enunciated in Euclid’s elements. Hilbert had already applied the new methods to the axiomatic theory of geometry, achieving standards of rigour which for the first time surpassed those found in Euclid’s geometry. This method envisages a special language for each branch of mathematics in which the primitives were “implicitly” defined by systems of axioms.

This is a substantial break with the ideas of Frege, who was very fussy about definitions. It is essential in the development of mathematics to pay close attention to definitions, since making use of a definition which
is incoherent invites paradox and invalidates all proofs which depend on those definitions. Risks are also associated with undertaking definitions on a piecemeal basis (defining a function first over one domain, and then over some other domain), because of the possibility that these partial definitions might conflict, or that the partiality itself might result in unsound reasoning (Frege had no methods for reasoning about partial functions).

8.3 On the Need for Synthetic Propositions

Carnap’s pluralism had several sources. One was his metaphysical agnosticism, which made him disinclined to reject languages which embodied objectionable ontologies. A second was the gradual emergence after Frege’s Begriffsschrift, first of direct alternatives (serving essentially the same purpose) such as Russell’s Theory of Types and then of a diversity of formal languages serving a variety of purposes, such as modal logics. The third was the desire to have formal languages in which synthetic propositions could be expressed.

We will consider language planning here without addressing the last of these three concerns, confining ourselves to the problem in relation to languages in which only analytic propositions can be expressed.

[There is a problem of order here. I had material in the chapter on the architecture of knowledge which explains why and how we can do without formalization of synthetic propositions. I have moved it here, but it probably doesn’t fit yet.]

To give a relatively concrete starting point to an otherwise highly abstract process I begin by describing an architecture which meets many of the more basic requirements, and in terms of which further requirements might be made intelligible. This is an architecture for the support of interactive proof in a logical foundation system. In many respects this architecture answers to the needs of Carnap’s project, but falls short of what Leibniz envisaged. Before describing that architecture we describe the idea of a logical foundation system for mathematics.

8.4 Logical Foundation Systems

In describing the kind of architecture here proposed, it is useful to begin by describing certain kinds of system which have been already in use and which provide a starting point from which the proposed system may be regarded as a further development.

The kind of system I have in mind here is an “interactive theorem prover” for a “logical foundation system”. In this section I describe the idea of a logical foundation system. In the next, the structure of a typical interactive theorem prover working with a logical foundation system, and then the proposed architecture is approached as a transformation to that starting point.

The idea of a logical foundation system begins with Frege, and refers to the kind of logical system which is needed for his logicist project of reducing mathematics to formal logic.

A logical foundation system is a formal language, preferably with a well defined semantics, and with a deductive system, which has sufficient semantic expressiveness and deductive strength that in it the concepts of mathematics may be defined and the theorems of mathematics derived.

From the point of view of establishing Frege’s logicist thesis the system has to be in some sense “purely logical”, and many contemporary philosophers doubt that any such system can serve as a foundation for mathematics. From our present point of view however, this aspect is inessential, it is not necessary for us to enter here into the question of what is a logical truth. The important part is that the system, whether purely logical or not, be suitable for defining mathematical concepts and for deriving mathematical theorems from those definitions.

In speaking here of “definition” we do not intend definitions as mere syntactic abbreviation, but as the introduction of new constants into the language satisfying certain defining conditions. This is not to include implicit definition by arbitrary axiomatic extension as advocated in Hilbert’s modern conception of axiomatic method, though the primitive constants of the system will effectively be defined in that manner. It is required that the development of mathematics be possible by definitions which are conservative extensions of the logical system, and for which it is possible for the system to reliably confirm that the proposed extension is conservative, and hence that the extended system will be consistent if the system being extended was.

The axiomatic set theory ZFC is a logical foundation system in the terms described above (though the means of conservative extension are not normally regarded as part of ZFC itself). Higher order logic as formulated by Alonzo Church in his Simple Theory of Types (STT)
is another foundation system which has proved popular for applications of formal mathematics in Computer Science, because of its successful implementation in a number of interactive theorem provers.

8.5 Interactive Theorem Provers

There has been continued research on the use of computers for proving theorems since shortly after the invention of digital computers. Software for theorem proving can be divided into two kinds which are called "automatic theorem provers" and "interactive theorem provers" respectively.

An automatic theorem prover is one which is supplied a description of the proof required, and searches for a proof without further involvement of its user. This might be a theorem prover for first order logic which is supplied with a goal to be proven and some premises from which to prove it and is then expected to come up with a proof in first order logic. Generally, the user would provide additional information influencing the way in which the prover undertakes the search for the proof, but the search would not itself involve any further interaction with the user.

From the point of view of general mathematics, and especially from the point of view of applications in information systems engineering (say in software or hardware verification), an automatic prover may not be suitable because of the complexity of the context in which the proof must be conducted, or simply because the required theorem itself may be so complex that no completely automatic proof is feasible within the current state of the art.

For these situations the use of an "interactive theorem prover" may be preferable. Such a system presumes that overall control of the construction of the proof is in the hands of its user, who will himself determine the gross structure of the proof in which the computer will assist by filling in detail. Interactive theorem provers usually incorporate automatic proof techniques, and may interface with automatic provers to obtain the kinds of proof which are within their scope. They increasingly act as integration platforms for diverse more narrowly focussed proof automation facilities permitting a single problem to be solved by an appropriate mixture of methods.

The fact that the user has control over the proof architecture means that problems on a greater scale of complexity can be tackled by an interactive theorem prover, typically limited only by the amount of human effort needed to direct the proof and progressively reduce the problem to portions modest enough to be undertaken automatically.

Because of the complexity of the specifications in the context of which proofs with an interactive theorem prover, whether these be simply an aggregation of background mathematical concepts relevant to the particular branch of mathematics in which the advances are to be made, or the specification of complex mathematical models of computer software or hardware, an interactive theorem prover must be able to organized a non-trivial collection of interdependent constant specifications. With typical theorem provers for variants on Church's STT this will be as a collection of "theories".

A theory in this context is like a hybrid between the concept of theory in mathematical logic (which is a set of sentences in some logical language) and the idea of a module in a modular programming language. On the logical side it is the purpose of the theory to collect together the definitions (axioms conservatively extending the logical system) which determine some particular context in which reasoning will take place. The theory (as a data structure managed by an interactive theorem prover) may act as a repository for the theorems which are proven in the context of those definitions. The organization of definitions into a hierarchy of theories also serves to record interdependencies between constant definitions.

On the modularity side theories serve to control the scope of relevance of definitions, so that a constant is only in scope when all the constants upon which its definition depends are also in scope, and may control various aspects of the presentation of the formal material to the user, details of concrete syntax.

8.6 Theory Hierarchy as Knowledge Base

The interactive theorem prover and its theory hierarchy represents for us the status quo ante. That from which we seek to move forward in our successor to the projects of Leibniz and Carnap.

To see this we consider how this relates to the ideas of Leibniz and Carnap.

Digital computers sewn together already acts in some respects as the repository of all human knowledge, and thus serve a role similar to that for which Leibniz pro-
moted the development of academic journals and encyclopaedia. They do so in the rather trivial sense that all the very many academic journals which have appeared since the time of Leibniz are available online (for a price), but they also provide completely new ways of aggregating and disseminating scientific and other knowledge, which threaten to displace the now traditional academic journals. Beyond this computers store an ever growing volume of information about all aspects of our lives and the world around us.

However, this information is mainly stored as data, the significance of which is not itself systematically recorded. It is a database, not a knowledge base.

Though the theory hierarchy of an interactive theorem prover was conceived of and is implemented to support only the accumulation of relatively small numbers of definitions and theorems rather than data on a large scale, it is of a very general character. Logically it would suffice for all the purpose for which computers store and manipulate data.

To illustrate this in intermediate ground let us consider how these systems are in fact used, and make modest extrapolations in the direction of the ambitions of Leibniz and Carnap.

The predominant uses of interactive theorem provers are in the development of purely mathematical theories, or of the kinds of applied mathematics which appears in theoretical computer science and in the application of such theories to reasoning about various kinds of computer systems, be they software or hardware. The software side is less significant because software may plausibly be regarded as consisting of abstract entities, computer programs as analogous with mathematical functions. The hardware is significant because we here venture into reasoning about something concrete rather than abstract. We are here using logic and mathematics to reason about the behavior of large scale electronic artifacts.

This is significant for our conception of Carnap’s program. One of the particular problems which Carnap grasped in order to support the application of modern logical methods in science was the extension of those logical methods to languages which speak of the concrete world rather than simply to languages in which one could speak only of abstract entities. This was the spur to Carnap’s pluralism, to his interest in the problem of defining languages and their semantics, and in languages in which synthetic propositions could be expressed. He departed from the universalistic perspec-
Chapter 9
The Architecture of Knowledge

It is a thesis of this work that the advancement of information technology renders choice of analytic method, not only in philosophy but wherever analysis might prove useful, dependent upon the software available to support the method, and that the architectural design of such software should take place in the context of an explicit (if generic or pluralistic) conception of analytic method.

This is an interdependency which may usefully be considered at the very earliest stages and at the highest and most abstract levels in the development of method and of information or knowledge architecture. The interdependency is such that we may be tempted to identify a certain kind of architectural design with a certain kind of fundamental philosophy or meta-philosophy.

In this chapter I undertake an analysis of knowledge architectures based on ideas about knowledge which have been so far presented. It is not desirable that an architectural discussion in a book of philosophy enter into much specific detail, so the aim here will be the analysis of certain ideas about the structure of knowledge as represented in information systems and the interaction between such conceptions of structure with the kinds of functionality which the information systems might then support, the methods which they facilitate, and the directions of future development to which they are sympathetic.

Rather than attempting wholly to effect the kind of integration between architectural design and constructive philosophical analysis, I will present these as two different perspectives upon a single enterprise, in this chapter the architectural design, in the next the philosophical perspective.

In this chapter the discussion will fall into two parts. An architecture is an abstract high level description. The analysis or evaluation of an architecture, must be undertaken against some prior conception of the aims which the architecture is intended to realize. In engineering terms these are high-level requirements. In philosophical terms, these requirements correspond to a delineation of the problem domain.

The first level at which positive philosophy departs from being purely analytic is in the choice of subject matter. A substantive statement about some practical matter, perhaps in politics or economics, may be clothed in a pure analysis, implicit in the choice of system to be studied. To promulgate ideas about how society might be organized, it would suffice to proceed by analysis, considering a class of realizations of the ideas and examining their relative merits.

This is the manner in which I proceed here. My interest is in certain approaches to the development of knowledge as a collaborative enterprise (which usually is) making effective use of our developing information and network capabilities. I begin the architectural discussion by setting out the domain of enquiry as a statement of requirements, and then proceed to consider and compare some of the ways in which those requirements might be met.

These two stages, statement of requirements, response to requirements, will not be monolithic. The requirements will be stated little by little. To each stage architectural responses are considered, and the requirements may then be augmented in the light of the analysis.

9.1 Requirements from Leibniz and Carnap

I have already identified the projects of Leibniz and Carnap as points of departure, so I begin with some
key features of those projects.

The principal elements of Leibniz’s project were:

- A Universal Language.
- An Encyclopaedia encompassing all scientific knowledge.
- A Calculus Ratiocinator for deciding truth.

Carnap’s project was more narrowly scoped, but shared the first two items recast in pluralistic terms.

Here we adopt all three, adjusting the statement in the direction of pluralism, and thinking of information technology.

Thus we are interested here in:

RA1 A system for the representation of propositional knowledge.

RA2 A substantial online knowledge base represented in that system.

RA3 Software supporting the further development and application of the knowledge.

Carnap attached considerable importance to the distinction between analytic and synthetic propositions. Though he did not acknowledge the influence of Hume, he agrees with Hume in characterizing essentially the same dichotomy, Hume’s fork, in three distinct ways. Carnap sought to adapt methods similar to those of Frege and Russell in the formal derivation of analytic propositions to languages in which synthetic propositions could be expressed and used in formal derivations.

Carnap’s approach to the meta-theory of such empirical languages is not from our point of view entirely satisfactory. The approach envisaged here to the connection of our languages with the empirical world is entirely different. Whereas

Formality and deduction also potentially enable new kinds of functionality to be realized. This is because the search processes involved in finding proofs of logical conjectures can serve to discover witnesses for existential claims, and hence solutions to design problems. The ability to demonstrate compliance of such a solution underpins and sanitizes the application of exotic and possibly unreliable methods during the search for a solution, hence allowing solutions to be discovered which might never be found by less exotic algorithms.

9.2 Epistemic Retreat

The positivistic idea of epistemic retreat influence the requirement in various ways.

Firstly we distinguish between analytic and synthetic judgements, and between formal and informal claims. The system is primarily concerned with the formal side.

As in Hume, we accept analytic propositions as exhausting those which can be known with certainty. In fact we go one further, taking empirical claims to be at best approximations, best thought of and formally represented as models of aspects of reality. As such they should be assessed or affirmed not as true or false but in more complex and informative terms. Thus, what we affirm of a theory is not its truth but its applicability to certain aspects of reality and the accuracy and reliability with which it models those aspects under various circumstances.

Formally, we do not assert an empirical claim in connection with such a model, we instead formulate the theory as an abstract model, and on the basis of this definition we can then undertake theoretical elaboration of the theory, draw consequences in relation to its application in hypothetical situations, and formally evaluate the theory against experimental data presented in the terms of the abstract model.

- RE1 Only analytic propositions are formally asserted.
- RE2 Empirical theories are represented as abstract models.
- RE3 Judgements are qualified by a measure of assurance.
Chapter 10

Metaphysical Positivism

*Metaphysical Positivism* is the name I have given to the theoretical aspects of Positive Philosophy, which is itself of broader scope.

The reason for naming a philosophical manner in this case is primarily for convenience of reference. Metaphysical positivism provides certain background materials in the context of which positive philosophy must be understood. Positive philosophy itself is simply philosophy conducted in the context of that background. The boundary between the two is not sharp but approximates to the traditional division between theoretical philosophy and practical philosophy.

In theoretical philosophy we consider those aspects of philosophy which have greatest impact on philosophical method, viz. metaphysics, logic, epistemology, and certain aspects of scientific method.

Metaphysical positivism is so called firstly because of its close connection with *logical positivism*, particularly with the philosophy of Rudolf Carnap, and because of the single most striking difference between it and logical positivism, which is in its use of the term *metaphysics*. It may therefore be helpful to begin with some remarks about the similarities and differences between metaphysical positivism and its predecessor.

### 10.1 First Base

*Metaphysical Positivism* is principally concerned with the *foundations of knowledge*. It therefore places *epistemology*, the theory of knowledge, to a central place in theoretical philosophy.

If we seek to build an enduring structure, it is best to build on a solid foundation. Critics of foundationalisms have taken foundationalism as demanding that such a foundation be immune to doubt, that it be, as a foundation for knowledge, absolutely solid. The foundationalism of Metaphysical positivism is not predicated on the existence of such foundations. It is rather the more pragmatic aim, given that we must start somewhere, to find the best place to start from. A foundation is therefore not to be absolutely solid, but just solid enough; fit for purpose.

When I speak here of a foundation as a place to start, this should not be taken too strictly. When we build a house, we begin with the foundations. The construction of the foundations may take perhaps one third of the time required for completing the house. In this case, the foundation is not something which we simply identify and use as a starting point. It is a stage in the construction, after which the character of the enterprise changes.

The foundation provides something on which a house can be built. The reason why the house stands securely is because it is built on a solid foundation. This is not the reason why a foundation is solid. A foundation is not evaluated in the same way as the house. Sometimes a foundation is solid because it consists of concrete laid on solid rock. Sometimes a foundation is solid because it is a concrete raft laid on something a much less rigid, perhaps clay. Sometimes a foundation is solid because it is made by driving piles into ground which is rather soft.

Ultimately, in these cases, the criteria are pragmatic and based on experience. The best foundation for a building is chosen taking into account the nature of the land on which it is to be build, the kind of building it is to support, and a great deal of experience and scientific knowledge of how different kinds of foundation will behave in these circumstances.

The foundationalism of Metaphysical Positivism is similarly pragmatic. It consists of ideas about what
ways of establishing, evaluating and applying different kinds of knowledge have proven effective, and on what methods we may expect to be effective as information technology and other factors transform the way we work with knowledge in the future.

The foundationalism of Metaphysical Positivism is thus self-consciously futuristic, it is oriented towards ways of working which will make the most of future advances in information engineering. The pragmatic aspect brings with it an epistemological pluralism. The foundationalisms I here espouse are not offered to the exclusion of other approaches.

There are three major stages in the foundations, closely connected with Hume’s two forks. The stages are addressed in sequence, each building on the earlier foundations.

An important part of the foundations proposed is simply conceptual. It is in the adoption of certain concepts with relatively definite meanings. The first concepts to consider are those which distinguish the three kinds of foundation.

The most fundamental of these comes from Hume’s fork, and the principal concept which we associate with this dichotomy which Hume identified is that of analyticity. Our first foundations are therefore foundations for analytic truth, and this is the main focus of the discussion here.

In relation to analytic truth, we do not advocate that any proposition be regarded in an unqualified way. The suggestion is that our system involves only the expression of opinions on analyticity, and that such opinions are in general expressed as based on certain other opinions. In very many cases these will be very solid opinions. Often the opinion will be the “opinion” of an interactive proof tool which has constructed and checked a formal proof of the proposition and made use of no other opinion.

The second kind of foundationalism concerns synthetic propositions, the other side of Hume’s first fork. Let us think of scientific laws as typical of this kind of knowledge. In respect of such laws I do not advocate that these should be considered in terms of truth or falsity. Experience tells us that scientific law are generally no more than approximations to “the truth”. Whether or not this is always the case, there are sufficiently many useful scientific laws which are known not to be strictly true that an epistemology which recognizes the merits of these is desirable.

Scientific laws are therefore considered as models of aspects of reality which are not regarded as either true or false, but as more or less accurate and reliable models of various aspects of the real world. The construction of such models is a purely logical matter, and the theoretical aspects of science which consist in the analysis of such logical models are covered under the foundational proposals for analytic truth. The new epistemological problems which arise concern the relationship between these abstract theories and those aspects of the world which they model.

10.2 By Comparison with Logical Positivism

Here and throughout, whenever I speak of logical positivism this should be understood to refer specifically to the philosophy of Rudolf Carnap whenever it concerns a matter on which the logical positivists may not have been unanimous.

The headline contrast with logical positivism is in relation to the word metaphysics, which I use in a manner quite distinct from the way in which it is used by Carnap. The best known feature of Carnap’s philosophy is his repudiation of metaphysics, around which, it is easy to suppose, his entire philosophy revolves. Metaphysical positivism embraces metaphysics, but the kinds of metaphysics which are accepted are not the kinds which were rejected by Carnap.

Metaphysics for Carnap is construed in very specific ways, and rather more narrowly than is usual in the positivistic tradition. Positivism is usually associated with nominalism, and involves the denial that abstract entities exist. Carnap on the other hand, was an ontological pragmatist, it sufficed for him that reasoning abstract entities was convenient for science to justify their use. His paper *Empiricism, Semantics and Ontology* is an exposition of his liberal attitudes in these matters. The metaphysics which Carnap did reject fell primarily under two headings. The first is the synthetic *a priori*, the second heading covers claims which have no definite meaning.

The first of these categories in Carnap’s conception of metaphysics is void likewise in Metaphysical Positivism. It is so partly because of the definitions (which are adopted in metaphysical positivism) of the concepts, and partly as an adopted epistemological criterion. Of these, more later.

So far as those which fail to be synthetic because
they are meaningless, the position of metaphysical positivism is softer. It is in the nature of philosophy that it
uses or investigates concepts whose meaning may be uncertain, or difficult to articulate. Dogmatic scepticism
about meaning in various degrees is common among academic logicians and philosophers, and is not a fea-
ture of metaphysical positivism. In this context by *dogmatic* scepticism we mean the movement from in-
comprehension to rejection. Our position in relation to doubt about meaning is to reserve judgement. How-
ever, if the proposition in doubt were offered as synthetic a priori, in the sense in which these terms are
understood in metaphysical positivism, then a firmer rejection would be called for.

Since Kripke the rejection of the synthetic a priori has generally been supposed to have been refuted. How-
ever, it can be seen that insofar as the relevant arguments are sound, then they must relate to concepts
distinct from those adopted in metaphysical positivism (and distinct from these concepts as used by Carnap).
The first step in showing this is to note that Carnap *defines* necessity as in terms of analyticity.

Having seen the historical development of philosophical positivism, having reconsidered positivism in the
light of the principal criticisms which were levelled at its most recent manifestation in *Logical Positivism*, and
taking account of certain ideas on about how information technology may transform the nature of knowledge,
it is now time to draw these themes together in a concisely stated positivistic synthesis.

Metaphysical Positivism is a graduated, constructive scepticism. In describing it as sceptical the emphasis is
placed upon an open minded suspension of judgement.

This suspension is graduated, and does not deny apparent and sometimes quite radical differences in our
confidence of working hypotheses. The most fundamental of such differences are associated with that between
logical and empirical knowledge associated with the analytic/synthetic distinction, and this leads to quite dif-
ferent ways of evaluating and affirming analytic and synthetic hypotheses. The constructive side of this
scepticism leads us into an epistemology which is coupled with architectural principles for the the future ex-
pansion of our knowledge in the context of a globally shared information infrastructure.

**10.3 Principal Features**

Metaphysical positivism is primarily concerned with analytic method, and with the conceptual framework
in which such methods can be articulated and evaluated and applied.

It is both linguistically and methodologically pluralistic, as in Carnap’s pluralism language is adopted on
the basis of pragmatic considerations. We are however aware, as Carnap was, that choice of vocabulary is im-
portant, and there is no suggestion that these choices are arbitrary.
Chapter 11
Digging Deeper

11.1 Foundations for Knowledge

How are we to judge claims to scientific knowledge?

In metaphysical philosophy and natural philosophy there have been ideas on this topic which may be called foundational. In this chapter we present some ideas along these lines.

Principally this concerns foundations for logical and metaphysical knowledge, knowledge *a priori*, but I will touch upon foundational aspects of empirical knowledge.

Before entering into a positive account of foundations I want to say a few words about the role which such foundations are intended here to fulfill.

It may be useful to draw an analogy with the use of the term ‘foundation’ in the construction of buildings. In the context a foundation provides a base solid enough for the construction of the desired building, so that the building will stand firm and will survive the stresses to which it may reasonably be expected to be subject.

For this a foundation does not need to be *absolutely* solid.

The construction of a foundation does not itself proceed in the same way as that for the building. One does not, in order to obtain a solid foundation, seek a yet more solid foundation on which to build the foundation (though sometimes bedrock serves this function). There does not arise in this way, an unsolvable problem of regress in the foundations of buildings.

There are two reasons

11.2 Logical Foundations

In keeping with the positivist tendency to which it belongs, our account of metaphysical positivism has been concerned primarily with underpinning and articulating methods and tools suitable not only for rigorous philosophical reasoning but for application in science and engineering.

In Aristotle’s conception of first philosophy, utility is regarded with some scorn, and the insistence of positivists that philosophy should facilitate positive science leads to the idea that positivism is an *anti-philosophical* philosophy (which is consistent with seeing it as continuous with academic and pyrrhonean scepticism).

Metaphysics is the name by which those topics at the apex of philosophy as conceived by Aristotle is now known, and the name “metaphysical positivism” may therefore be read as hinting that the pragmatic orientation of our positivism does not involve a rejection of those more remote regions of philosophy whose connection with life seems most tenuous.

Nevertheless, in metaphysical positivism, locating a place for metaphysical investigation is not easy. Two kinds of defect which may be found in metaphysical (and other controversy) from a positivistic standpoint are meaningless claims and purely verbal disputes. It is characteristic of positivist to reject metaphysics as meaningless, and in metaphysical positivism I retain a concern for precision and clarity in language, which motivates some of the deeper concerns which we address here. However, it is the business of philosophy to address problems whose articulation is difficult, and that one philosopher does not find a conjecture or a definition meaningful does not suffice to establish that it is not.

In keeping with the graduated scepticism in metaphysical positivism meaningfulness is not taken to be an all or nothing affair. Languages (or idiolects) may be compared on two related kinds of scale. First they may be compared according to their expressiveness. A
11.2. LOGICAL FOUNDATIONS

language A is as expressive as language B if everything which can be said in language B can also be said in language A. To compare precision of definiteness of a language we have to consider languages as having multiple possible meanings or interpretations and then compare the range of interpretations of two languages.

An easy and fundamental illustration of this kind of comparison may be found in axiomatic set theory. If we consider a specific theory, say ZFC, the axioms of the theory provide an implicit definition of the concept of set which is the subject matter of the theory. The truth conditions of sentences of ZFC can be made very definite by stipulating that a sentence is true in ZFC iff it is true in every model of the axioms. Truth will then correspond to provability, in consequence of the completeness of first order logic. It is also reasonable in this domain to take the meaning to be the truth conditions, so that ZFC becomes as definite in its meaning as first order logic is. Unfortunately when we look at the intended applications of set theory, of which the first is to the theory of arithmetic, we find that this conception of the meaning of set theory is unsatisfactory. The normal procedure in reasoning about arithmetic in set theory is to define the natural numbers as some convenient countably infinite set of representatives. The most popular has been the scheme for representation of ordinal numbers under which the natural number zero is represented by the empty set, and every other natural number is represented by the set of its predecessors, i.e. the set of all natural numbers which are less than that number.

Using this definition, together with definitions of the usual arithmetic operations over these representatives, we can derive the usual theorems of arithmetic in ZFC. More true theorems of arithmetic are provable in this way than is possible in the usual direct axiomatization of arithmetic in first order logic (known as PA, for Peano Arithmetic). However, it is known, as a result related to the incompleteness results proved by Kurt Gödel that not all the truths of arithmetic are provable in this way.

However, because of the completeness of first order logic, all the statements of arithmetic as expressed in the way indicated in ZFC which are true under that semantics. The arithmetic truths which are not provable in ZFC, are, under that semantics, with that manner of representation of numbers, not even true. We have failed to produce an adequate definition of the natural numbers.

This is not an avoidable defect in the Von Neumann representation of ordinals. It is easy to see that under the proposed semantics for ZFC the truths of set theory will be recursively enumerable, and therefore any decidable subset of those truths (the truths of set theory which happen to correspond to sentences of first order arithmetic) will also be effectively enumerable, whatever definition of natural number we start out with. But it is know that the truths of arithmetic are not recursively enumerable. It follows that the concept of natural number is not representable in ZFC under the given semantics.

Though the semantics is definite, it is not expressive.

We can make the semantics more expressive, at the cost of making it less definite, in the following way. The semantics we have been discussing involves acceptance of all models of ZFC, and its (semantic) incompleteness reflects the existence of models of ZFC in which the set defined to be the natural numbers is not what the definition is intended to give.

The definition of the natural numbers is intended to give a set whose members are all the sets obtainable from the empty set by repeated application of the successor function (the function s(x) = x+1). The idea “obtainable by repeated application” cannot be directly expressed in first order logic, so the definition instead is given in terms of closure under the successor function. A set is closed under the successor function if for every member of the set, its successor is also a member. The natural numbers are then defined as the intersection of all sets which contain the empty set and are closed under the successor function.

Unfortunately, if we take an interpretation in which the intended set of natural numbers does not exist, in which every set which contains all the natural numbers also contains some other set, then when you take the intersection you get a set which contains that other set. This is a model with non-standard natural numbers, and such an interpretation will not get the truths of arithmetic right. Because our semantics allows these non-standard models, as well as models in which arithmetic is standard, statements of arithmetic which are true in the standard models but which are violated in some non-standard model will come out under the semantics as false.

To get the semantics on the nose for arithmetic statements we need to eliminate these non-standard models from the semantics. We cannot do this by adding another axiom, because the required constraint on the
models is not expressible in first order logic. But we can add an informal stipulation to the semantics. We can specify the truth condition for sentences in ZFC as truth in all models of ZFC with standard natural numbers.

We now have a version of ZFC which is more expressive than the previous one. One in which the natural numbers really are definable (though not in the sense of this term which is used by mathematical logicians) and in which the sentences of arithmetic have the correct truth values. Though the semantics of this language are defined in part informally, the language can now be used to define the semantics of other languages in a formal way (relatively), whose semantics would not be definable in first order logic.

In this way we can define variants of the language of first order set theory which have progressively greater expressive power. This is done by using informal constraints on the class of intended interpretations of the theory. Such informal constraints can be placed in order of strength and the stronger the constraint is the more expressive the resulting language will be.

Beyond the constraint to models with standard natural numbers the following stronger constraints can be applied:

- Constraining interpretations to be well-founded.
- Requiring full power-sets.

Constraining interpretations to be well-founded is strictly stronger than requiring standard natural numbers. It is as strong because in any model in which the set of natural numbers is non-standard it is not well-founded. It is strictly stronger because the same consideration applies to all limit ordinals (all transfinite numbers). In the absence of well-foundedness we can have models which give standard natural numbers but have a non-standard ordinal somewhere higher in the hierarchy. Well-founded models not only have standard arithmetic, but they also have standard ordinals all the way up. So under this semantics, but not under the previous one, the ordinals are definable.

Yet greater strength is obtained by requiring the full power set. All models of ZFC are closed under the formation of power sets. That means, that for every set in the domain of discourse the collection of its subsets (in the domain of discourse) is also a set in the domain of discourse. However, it is not the case that the power set is the same in every model, since not all subsets of the set are bound to be in the domain of discourse. The power set will include all the subsets which exist, but there may be subsets which don’t exist (in this particular model). The constraint to full power sets eliminates any model in which there are missing subsets.

If we define truth in ZFC as truth in every model of ZFC in which the power set is full, then we get another language which is again strictly more expressive than the ones we have previously considered. Why is this? To understand this it is helpful to address the question why there exist non-well-founded models of ZFC.

There is an axiom in ZFC which is intended to assert that all sets are well-founded. This is called the axiom of regularity. Its intended effect is to deny that there are any infinite descending chains in the membership relationship, but this, like the obvious informal definition of the natural numbers, cannot be directly stated informally. The axiom of regularity states instead that every set has a minimal element. A minimal element of a set A is a member B of A which contains no member of A, such that A intersection B is empty.

This definition does ensure well foundedness if there are enough sets in the domain of discourse, which is the case if we have full power sets. Otherwise it does not, and there are non-well founded models of ZFC despite it having an axiom intended to deny their existence. Well foundedness is not expressible in first order logic, and so cannot be fully incorporated into the axioms, which is why the previous semantics relies on an informal constraint to well-foundedness.

If instead of requiring well-foundedness we stipulate the semantics in terms of models with full power sets, then since these are all well founded the resulting semantics is as expressive as the semantics based on arbitrary well-founded models.

That the semantics is strictly stronger can be seen from consideration of cardinal numbers. Cardinal numbers may be represented in ZFC as initial ordinals. An initial ordinal is an ordinal which has a greater cardinality than any previous ordinal. Two sets have the same cardinality if there exists a bijection between their elements. Such a bijection pairs up the elements in the two sets in a one-one manner so that they can be seen to have the same size. A difficulty with this definition of cardinality arises from the existence of models in which not all subsets of every set are present, in which we do not have full power sets. This is because the non-existence of a bijection might arise not because the two sets really are of different size, but because the bijection...
between their elements is just not in the domain of the model. The effect of this is that not all models of ZFC agree about which ordinals are initial, and consequently they do not agree about cardinal arithmetic.

Constraining the truth conditions to involve truth only in models with full power sets eliminates this source of disagreement between intended models of ZFC about cardinal arithmetic. Under this semantics but not under any of the previous semantics we can define the notion of cardinal number, so it gives a strictly more expressive language.

11.3 Empirical Foundations
Bibliography


[Car63a] Rudolf Carnap. Intellectual autobiography. In Schilpp [Car63b].


[Rus18] Bertrand Russell. The philosophy of logical atomism. 1918.


List of Tables

3.1 Development of the Analytic/Synthetic distinction .................. 21
6.1 Rigour, Scepticism, Positivism ......................................... 35
Positions

3.1 the world of forms and world of appearances .......................................................... 17
3.2 the a priori (knowledge) and a posteriori (opinion) .................................................. 17
3.3 essential and accidental predication .......................................................................... 17
3.4 definition and essence .............................................................................................. 17
Index

a priori, 20
analysis, 22
analytic, 22
analytical engine, 27
appearances, 17
axiomatic method, 16

Bacon, 17
Begriffsschrift, 27
calculus ratiocinator, 7, 29
Carnap, Rudolf, 13
Church, 29
Church, Alonzo, 40
Comte, Auguste, 34

Empiricism, 17
Empiricus, Sextus, 33
equipollence, 33
Euclid, 27

Foundations
of knowledge, 45

Heraclitus, 17
Hilbert, David, 27, 28
Hobbes, 17
Hume
enquiry, 14
treatise, 14

ideal forms, 17

Kant, 22
Kleene, 29
Kolakowski, 34
Kripke, Saul, 6
Kronecker, Leopold, 28
Kurt Gödel, 49

Leibniz, 13, 29

Locke, 17
logical foundation system, 40

meta-mathematics, 27
Metaphysical Positivism Metaphysical Positivism, 45
Moore, G.E., 24

Parmenides, 17
positivism, 34
Post, 29
Principia Mathematica, 21

Russell, 18
Russell, Bertrand, 24, 27

Simple Theory of Types, 40
STT, 40
sufficient reason, 18
synthesis, 22
synthetic, 22

Thales, 26
the LCF paradigm, 29
Turing, 29

universal characteristic, 7, 27

Von Neumann, 49

Whitehead, A.N., 24
Wittgenstein, 18
Wittgenstein, Ludwig, 6

Zeno of Elea, 17
ZFC, 40, 49