CHAPTER 1

A Brief History of Artificial Intelligence

What is all knowledge too but recorded experience, and a product of history; of which, therefore, reasoning and belief, no less than action and passion, are essential materials?

—Thomas Carlyle, *Critical and Miscellaneous Essays*

History is Philosophy from Examples.

—Dionysius, *Ars Rhetorica*

Science is built upon facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house.

—Henri Poincaré, *Science and Hypothesis*

You seek for knowledge and wisdom as I once did; and I ardently hope that the gratification of your wishes may not be a serpent to sting you, as mine has been.

—Mary Shelley, Frankenstein

1.1 Introduction

Although Artificial Intelligence is one of the newest fields of intellectual research, its foundations began thousands of years ago. In studying Artificial Intelligence, it is useful to have an understanding of the background of a number of other subjects, primarily philosophy, linguistics, psychology, and biology.

This chapter will present a selected history of the thinking and research that led up to the present state of what we now call Artificial Intelligence.

In this chapter, we will look at the contributions made by philosophy, linguistics, psychology, and biology to Artificial Intelligence. We will also look at the difference between the claims made by proponents of weak AI (AI is a commonly used abbreviation for Artificial Intelligence) compared with those who support strong AI, as well as look at the difference between strong methods and weak methods in Artificial Intelligence.

We will begin by looking at Artificial Intelligence itself and trying to find a definition for the subject.

1.2 What Is Artificial Intelligence?

Perhaps a better starting point would be to ask, "What is intelligence?" This is a complex question with no well-defined answer that has puzzled biologists, psychologists, and philosophers for centuries. In Chapter 13 we pose a similar question when we ask, "What is life?" in order to help us understand what Artificial Life, a branch of Artificial Intelligence, is.

One could certainly define intelligence by the properties it exhibits: an ability to deal with new situations; the ability to solve problems, to answer questions, to devise plans, and so on. It is perhaps harder to define the difference between the intelligence exhibited by humans and that exhibited by dolphins or apes.

For now we will confine ourselves, then, to the somewhat simpler question that is posed by the title of this section: What Is Artificial Intelligence?

A simple definition might be as follows:

Artificial intelligence is the study of systems that act in a way that to any observer would appear to be intelligent.

This definition is fine, but in fact it does not cover the whole of Artificial Intelligence. In many cases, Artificial Intelligence techniques are used to solve relatively simple problems or complex problems that are internal to more complex systems. For example, the search techniques described in Chapter 4 are rarely used to provide a robot with the ability to find its way out of a maze, but are frequently used for much more prosaic problems.

This may lead us to another definition of Artificial Intelligence, as follows:

Artificial Intelligence involves using methods based on the intelligent behavior of humans and other animals to solve complex problems.

Hence, in Chapter 20, we look at systems that are able to "understand" human speech, or at least are able to extract some meaning from human utterances, and carry out actions based on those utterances. Such systems may not be designed to behave in an intelligent way, but simply to provide some useful function. The methods they use, however, are based on the intelligent behavior of humans.

This distinction is brought into sharper contrast when we look at the difference between so-called **strong AI** and **weak AI**.

The followers of strong AI believe that by giving a computer program sufficient processing power, and by providing it with enough intelligence, one can create a computer that can literally think and is conscious in the same way that a human is conscious.

Many philosophers and Artificial Intelligence researchers consider this view to be false, and even ludicrous. The possibility of creating a robot with emotions and real consciousness is one that is often explored in the realms of science fiction but is rarely considered to be a goal of Artificial Intelligence.

Weak AI, in contrast, is simply the view that intelligent behavior can be modeled and used by computers to solve complex problems. This point of view argues that just because a computer behaves intelligently does not prove that it is actually intelligent in the way that a human is. We will examine this argument in more detail in Chapter 2, when we look at the Chinese Room thought experiment and the arguments around it.

1.3 Strong Methods and Weak Methods

We have discussed the difference between the claims of weak AI and strong AI. This difference is not to be confused with the difference between **strong methods** and **weak methods**.

Weak methods in Artificial Intelligence use systems such as logic, automated reasoning, and other general structures that can be applied to a wide range of problems but that do not necessarily incorporate any real knowledge about the world of the problem that is being solved.

In contrast, strong method problem solving depends on a system being given a great deal of knowledge about its world and the problems that it might encounter. Strong method problem solving depends on the weak methods because a system with knowledge is useless without some methodology for handling that knowledge.

Hence, the production systems we will examine in Chapter 9 are based on the weak method expert system shells but use strong method rules to encode their knowledge.

The earliest research in Artificial Intelligence focused on weak methods. Newell and Simon's General Problem Solver (GPS), which is discussed in Chapter 15, was an attempt to use weak methods to build a system that could solve a wide range of general problems. That this approach ultimately failed led to a realization that more was needed than simple representations and algorithms to make Artificial Intelligence work: knowledge was the key ingredient.

A great number of the subjects covered in this book are weak methods. This does not mean that they are not worth studying, or even that they are not useful. In many situations, weak methods are ideal for solving problems. However, the addition of knowledge is almost always essential to build systems that are able to deal intelligently with new problems; if our aim is to build systems that appear to behave intelligently, then strong methods are certainly essential.

1.4 From Aristotle to Babbage

In Chapter 7 of this book, we present the propositional and predicate logics. These systems for logical reasoning are based on the logic invented by Aristotle, a philosopher from ancient Greece, who lived from 384 to 322 B.C. and who studied under Plato during that time. The writings of Aristotle (on this and many other subjects) have formed the basis for a great deal of our modern scientific thinking.

From the point of view of Artificial Intelligence, the most interesting aspect of Aristotle's work is his study of logic. He invented the idea of the **syllogism**, which he defined as follows:

> "A discourse in which certain things having been stated, something else follows of necessity from their being so."

Aristotle's logic was developed and expanded on by later philosophers, mathematicians, and logicians. The first real steps in the study of logic after Aristotle took place in the 12th century, when Peter Abelard (who lived from 1079 to 1142 A.D.) wrote *Dialectica*, a treatise on logic. In the following centuries, more work was carried out, but the greatest developments were made in the last few centuries.

In the late 17th to early 18th centuries, Gottfried Leibniz, the German mathematician and philosopher who along with Isaac Newton had a part in the invention of the calculus used by mathematicians today, invented the idea of developing a formal mathematical language for reasoning. His universal language would allow us to express with great precision problems of all kinds, and then go about solving them. Leibniz did not succeed in creating this universal language, but his work provided the basis for the propositional and predicate logics that are so important to Artificial Intelligence research today.

In the 19th century, George Boole, an English mathematician, who lived from 1815 to 1864, developed Boolean algebra, the logical system we still use as part of propositional and predicate logics. Boolean algebra is widely used by electronics engineers in developing logical gates for silicon chips and is also used by computer scientists. Boolean algebra provides a language for expressing concepts such as "A is true" and "A is true but B is false."

Around the same time that Boole was inventing his algebra, Charles Babbage invented the world's first computer—the Analytic Engine. He didn't ever manage to build the computer, but his designs were later used to build a working model. The designs of computers in the 20th century didn't bear much resemblance to Babbage's computer, but they certainly owed a great deal to it.

Babbage's idea of a digital computer remained a dream until around the middle of the 20th century. By the 1950s, a number of working computers had been built. Unlike Babbage's mechanical engines, these computers were electronic. The very first electromechanical computers were soon replaced by computers based on vacuum tubes.

1.5 Alan Turing and the 1950s

One of the great figures in the history of Artificial Intelligence is Alan Turing. During World War II, Turing famously worked in Bletchley Park, helping to solve the Germans' codes. After the war, he began to work on the idea of the possibility of building a computer that could think. His paper published in 1950, *Computing Machinery & Intelligence*, was one of the first papers to be written on this subject.

The **Turing test** was designed by Turing as a way to judge the success or otherwise of an attempt to produce a thinking computer. More specifically, it was based on the idea that if a person who interrogated the computer could not tell if it was a human or a computer, then to all intents and purposes, Turing said, it is intelligent.

The test is designed as follows:

The interrogator is given access to two individuals, one of whom is a human and the other of whom is a computer. The interrogator can ask the two individuals questions, but cannot directly interact with them. Probably the questions are entered into a computer via a keyboard, and the responses appear on the computer screen.

The human is intended to attempt to help the interrogator, but if the computer is really intelligent enough, it should be able to fool the interrogator into being uncertain about which is the computer and which is the human.

The human can give answers such as "I'm the human—the other one is the computer," but of course, so can the computer. The real way in which the human proves his or her humanity is by giving complex answers that a computer could not be expected to comprehend. Of course, the inventors of the truly intelligent computer program would have given their program the ability to anticipate all such complexities.

Turing's test has resulted in a number of computer programs (such as Weizenbaum's ELIZA, designed in 1965) that were designed to mimic human conversation. Of course, this in itself is not a particularly useful function, but the attempt has led to improvements in understanding of areas such as natural language processing. To date, no program has passed the Turing test, although cash prizes are regularly offered to the inventor of the first computer program to do so.

Later in the 1950s computer programs began to be developed that could play games such as checkers and chess (see Chapter 6), and also the first work was carried out into developing computer programs that could understand human language (Chapter 20).

A great deal of work at this stage was done in computer translation. It was, indeed, widely believed that computers could eventually be programmed to translate accurately from one human language to another. It has since been found that the task of machine translation is actually an extremely difficult one, and not one that has yet been completely solved. This subject is discussed in more detail in Chapter 20.

In 1956, the term **Artificial Intelligence** was first used by John McCarthy at a conference in Dartmouth College, in Hanover, New Hampshire.

In 1957, Newell and Simon invented the idea of the GPS, whose purpose was, as the name suggests, to solve almost any logical problem. The program used a methodology known as means ends analysis, which is based on the idea of determining what needs to be done and then working out a way to do it. This works well enough for simple problems, but AI researchers soon realized that this kind of method could not be applied in such a general way—the GPS could solve some fairly specific problems for which it was ideally suited, but its name was really a misnomer.

At this time there was a great deal of optimism about Artificial Intelligence. Predictions that with hindsight appear rash were widespread. Many commentators were predicting that it would be only a few years before computers could be designed that would be at least as intelligent as real human beings and able to perform such tasks as beating the world champion at chess, translating from Russian into English, and navigating a car through a busy street. Some success has been made in the past 50 years with these problems and other similar ones, but no one has yet designed a computer that anyone would describe reasonably as being intelligent.

In 1958, McCarthy invented the LISP programming language, which is still widely used today in Artificial Intelligence research.

1.6 The 1960s to the 1990s

Since the 1950s, a great deal of the original optimism has gone out of Artificial Intelligence and has been replaced with a degree of realism.

The aim of the study of Artificial Intelligence is no longer to create a robot as intelligent as a human, but rather to use algorithms, heuristics, and methodologies based on the ways in which the human brain solves problems. Hence, systems have been designed such as Thomas Evans' **Analogy** and Melanie Mitchell's **Copycat Architecture**, which were designed to be able to solve problems that involve analogies. Mitchell's Copycat, for example, can solve problems such as "ABC is to CBA as DEF is to ???."

The ability to solve problems of this kind does not represent intelligence, but the development of systems that can solve such problems is the mainstay of Artificial Intelligence research and arguably an extremely useful step along the way to producing more and more useful computer software systems.

In Chapter 2, we will discuss the subject of whether a computer program can really be "intelligent."

In the most recent decades, the study of Artificial Intelligence has flourished. Areas of particular importance include the following:

- machine learning
- multi-agent systems
- artificial life
- computer vision
- planning
- playing games (chess in particular)

In Chapter 2, we will look at the prevalence of Artificial Intelligence in the world today. This prevalence has more than justified the work of the past 50 years.

1.7 Philosophy

The philosophy of great thinkers, from Plato to Descartes and to Daniel Dennett, has had a great deal of influence on the modern study of Artificial Intelligence.

The influence of Aristotle has already been mentioned, but it has been argued (Dreyfus, 1972) that the history of Artificial Intelligence begins when Plato wrote that his teacher Socrates said, "I want to know what is characteristic of piety which makes all actions pious. . . that I may have it to turn to, and to use as a standard whereby to judge your actions and those of other men."

Socrates was claiming that an algorithm could be defined that described the behavior of humans and determined whether a person's behavior was good or bad.

This leads us to a fundamental question that has been asked by philosophers and students of Artificial Intelligence for many years: Is there more to the mind than simply a collection of neurons? Or, to put it another way, if each neuron in the human brain was replaced by an equivalent computational device, would the resultant be the same person? Would it indeed be capable of intelligent thought?

This kind of question is regularly debated by modern philosophers such as Daniel Dennett, and while the answer is far from clear, it is an instructive debate to follow, and its implications for Artificial Intelligence are enormous.

In the 17th century, the great philosopher René Descartes was a strong believer in **dualism**, the idea that the universe consists of two entirely separate things: mind and matter. Descartes's view was that the mind (or soul) was entirely separate from the physical body and not constrained by it in any way.

Importantly, Descartes did not believe that this dualism extended to animals. In other words, in his view a cat or a dog is simply a machine: a highly complex machine, but a machine nonetheless. This view gives hope to the proponents of Artificial Intelligence who believe that by simply putting enough computing power together and programming it in the correct way, a machine could be made to behave in the same way as an animal, or even a human being.

1.8 Linguistics

The study of human language has a vital role to play in Artificial Intelligence. As is discussed in some detail in Chapter 20, compared with computer languages such as Java and LISP, human languages are extraordinarily complex and are full of pitfalls that almost seem designed to trap anyone (human or computer) inexperienced in the use of the language.

This complexity, combined with a sense of optimism, may well have been part of the reason that natural language processing was such a popular research area in the early days of Artificial Intelligence.

Some of the optimism surrounding Natural Language Processing came from the writings of Noam Chomsky, who in the 1950s proposed his theory of Syntactic Structures, which was a formal theory of the structure of human language. His theory also attempted to provide a structure for human knowledge, based on the knowledge of language.

This idea of knowledge representation is at the very core of Artificial Intelligence and is a recurring theme throughout this book.

Almost all of the techniques described in this book depend on a formal method of representation for knowledge that enables a computer to use information from the world, or concerning the problems it is to solve, without necessarily needing to understand that knowledge.

There is a close relationship between linguistics and Artificial Intelligence, and the two fields join together in the study of natural language processing, which is discussed in some detail in Chapter 20.

1.9 Human Psychology and Biology

Some of the techniques, such as search algorithms, described in this book do not clearly map onto any specific biological or psychological function of human beings. On the other hand, many of them do. For example, McCulloch and Pitts's electronic neurons, which are used today to build neural networks, are directly based on the way in which neurons in the human brain function.

In a similar way, much research in Artificial Intelligence has been related to **cognitive psychology**, which is based on the idea that the human brain uses knowledge or information that it is capable of processing in order to solve problems, make decisions, draw conclusions, and carry out other intelligent acts.

This form of psychology was in contrast to behaviorism, which prevailed for much of the first half of the 20th century. Behaviorism relates behavior directly to stimuli, without taking into account knowledge or information that might be contained in the brain. This is the kind of psychology that Pavlov was demonstrating in his famous experiment with dogs.

Psychology is certainly useful to the study of Artificial Intelligence in one respect: it helps to answer the important question, "What is intelligence?" As we have seen already, this is a difficult question to answer, but in studying it, psychologists give us a great deal of information that is useful in forming the ideas behind Artificial Intelligence.

1.10 AI Programming Languages

A number of programming languages exist that are used to build Artificial Intelligence systems. General programming languages such as C++ and Java are often used because these are the languages with which most computer scientists have experience. There also exist two programming languages that have features that make them particularly useful for programming Artificial Intelligence projects—PROLOG and LISP.

We will now provide a brief overview of these two languages and explain how they are used in Artificial Intelligence research. Of course, a number of other programming languages exist that are also widely used for Artificial Intelligence, but we will focus on PROLOG and LISP because these are certainly the most widely used and the ones on which there is the widest range of relevant literature.

1.10.1 PROLOG

PROLOG (PROgramming in LOGic) is a language designed to enable programmers to build a database of **facts** and **rules**, and then to have the system answer questions by a process of logical deduction using the facts and rules in the database.

Facts entered into a PROLOG database might look as follows:

```
tasty (cheese).
made_from (cheese, milk).
contains (milk, calcium).
```

These facts can be expressed as the following English statements:

Cheese is tasty.

Cheese is made from milk.

Milk contains calcium.

We can also specify rules in a similar way, which express relationships between objects and also provide the instructions that the PROLOG theorem prover will use to answer queries. The following is an example of a rule in PROLOG:

```
contains (X, Y) :- made_from (X, Z), contains (Z, Y).
```

This rule is made up of two main parts, separated by the symbol ":-".

The rule thus takes the form:

```
B :- A
```

which means "if A is true, then B is true," or "A implies B."

Hence, the rule given above can be translated as "If X is made from Z and Z contains Y then X contains Y."

In Chapters 7, 8, and 9, we make a great deal of use of rules of this kind.

Having entered the three facts and one rule given above, the user might want to ask the system a question:

```
?- contains (cheese, calcium).
```

Using a process known as resolution (which is described in detail in Chapter 8), the PROLOG system is able to use the rule and the facts to determine that because cheese is made from milk, and because milk contains calcium, therefore cheese does contain calcium. It thus responds:

```
yes
```

It would also be possible to ask the system to name everything that contains calcium:

```
?- contains (X, calcium)
```

The system will use the same rules and facts to deduce that milk and cheese both contain calcium, and so will respond:

```
X=milk.
X=cheese.
```

This has been a very simple example, but it should serve to illustrate how PROLOG works. Far more complex databases of facts and rules are routinely built using PROLOG, and in some cases simple databases are built that are able to solve complex mathematical problems.

PROLOG is not an efficient programming language, and so for many problems a language such as C++ would be more appropriate. In cases where logical deduction is all that is required, and the interactive nature of the PROLOG interface is suitable, then PROLOG is the clear choice. PROLOG provides a way for programmers to manipulate data in the form of rules and facts without needing to select algorithms or methodologies for handling those data.

1.10.2 LISP

LISP (LISt Programming) is a language that more closely resembles the imperative programming languages such as C++ and Pascal than does PROLOG. As its name suggests, LISP is based around handling of lists of data. A list in LISP is contained within brackets, such as:

[A B C]

This is a list of three items. LISP uses lists to represent data, but also to represent programs. Hence, a program in LISP can be treated as data. This introduces the possibility of writing self-modifying programs in LISP, and as we see in Chapter 13, it also allows us to use evolutionary techniques to "evolve" better LISP programs.

LISP is a far more complex language syntactically than PROLOG, and so we will not present any detail on its syntax here. It provides the usual kinds of mechanisms that other programming languages provide, such as assignment, looping, evaluating functions, and conditional control (if. . . then. . .). It also provides a great deal of list manipulation functions, such as car and cdr, which are used to return the first entry in a list and all the entries except for the first entry, respectively.

1.11 Chapter Summary

- Intelligence is difficult to define, and as a result Artificial Intelligence is also hard to define.
- One definition of Artificial Intelligence is:

  Artificial intelligence is the study of systems that act in a way that to any observer would appear to be intelligent.
- Proponents of strong AI believe that a computer that behaves in an intelligent way is capable of possessing mental states and, therefore, of being truly conscious and intelligent in the same way that humans are.
- Weak AI is a less controversial idea—that computers can be programmed to behave in intelligent ways in order to solve specific problems. This book is concerned with the methods of weak AI.
- Weak and strong AI are not to be confused with weak and strong methods.
- Weak methods are those that do not rely on any knowledge or understanding of the world and the problems being solved. Most of the techniques described in this book are weak methods.
- Strong methods are those that use knowledge about the world and about the problem being solved. The strong method approach is essential for solving many complex real world problems using Artificial Intelligence.

- In studying Artificial Intelligence, it is extremely useful to understand the background of philosophy, linguistics, biology, and psychology.
- Philosophers, from Plato and Aristotle to Searle and Dennett, have asked questions and provided opinions concerning the nature of intelligence and the ability to define it in a way that would enable us to program a computer with real intelligence.
- The 1950s were a time of great optimism in Artificial Intelligence and also a time of great progress in the field.
- Turing's test is a way to determine if a computer is truly intelligent, by seeing if it could fool a human in conversation into thinking that it too was human. It is widely believed today that even if a computer could pass the Turing test, it would still not truly be conscious or intelligent in the way that humans are.
- In 1956 the term Artificial Intelligence was coined by John McCarthy.
- Since the 1950s, the study of Artificial Intelligence has been flavored with a great deal more realism. The progress in recent years has been phenomenal.

1.12 Review Questions

1.1 What is intelligence?

1.2 What is Artificial Intelligence? What do you hope to learn by reading this book?

1.3 Is Artificial Intelligence a branch of computer science or an alternative to computer science?

1.4 Why is Artificial Intelligence a worthwhile subject to study?

1.5 Explain the difference between strong and weak methods in Artificial Intelligence. Explain how this dichotomy differs from the difference between strong and weak AI.

1.6 Why are PROLOG and LISP so well suited to Artificial Intelligence research? Do you think languages such as C++ and Java could also be used for such research?

1.7 What do you think led mankind to embark upon the study of Artificial Intelligence? Which fields of study particularly fed into it?

What human desires did the study of Artificial Intelligence seek to satisfy?

1.8 When did Artificial Intelligence first begin to be studied? Your answer should be more detailed than a simple date.

1.13 Further Reading

Crevier (1999) gives a fascinating history of the subject of Artificial Intelligence.

Throughout this book, details are given of other books that can be referenced to learn more about the material covered herein. The following books are general Artificial Intelligence texts that cover almost all of the topics covered by this book and also provide excellent introductions to the subject as a whole.

Each of these books takes a different approach to the material, and it is worth selecting the text that best fits your personal preferences in studying this subject.

For example, Russell and Norvig present the material in terms of intelligent agents. Winston explains his material with a great deal of examples but tends not to go into a great deal of detail, while Luger goes into greater depth, but with fewer examples. Schalkoff gives a good coverage of Artificial Intelligence using examples in PROLOG and LISP; it also therefore serves as a useful text in those languages.

Computation & Intelligence, edited by George Luger, contains a number of extremely important papers collected from the whole history of Artificial Intelligence. It includes papers by such pioneers of the subject as Alan Turing, Marvin Minsky, John McCarthy, Allen Newell, and Herbert Simon.

The Handbook of Artificial Intelligence, edited by A. Barr and E. Feigenbaum (1989 – William Kaufman)

The Essence of Artificial Intelligence, by Alison Cawsey (1998 – Prentice Hall)

Introduction to Artificial Intelligence, by Eugene Charniak and Drew McDermott (1985 – Addison Wesley; out of print)

The Computational Brain, by Patricia S. Churchland and Terrence J. Sejnowski (1992 – The MIT Press)

AI: The Tumultuous History of the Search for Artificial Intelligence, by Daniel Crevier (1999 – Basic Books)

Understanding Artificial Intelligence (Science Made Accessible), compiled by Sandy Fritz (2002 – Warner Books)

The Anatomy of Programming Languages, by Alice E. Fischer and Frances S. Grodzinsky (1993 – Prentice Hall)

Introduction to Artificial Intelligence, by Philip C. Jackson (1985 – Dover Publications)

AI Application Programming, by M. Tim Jones (2003 – Charles River Media)

Artificial Intelligence: Structures and Strategies for Complex Problem-Solving, by George F. Luger (2002 – Addison Wesley)

Computation & Intelligence: Collected Readings, edited by George F. Luger (1995 – The AAAI Press / The MIT Press)

Artificial Intelligence: A Guide to Intelligent Systems, by Michael Negnevitsky (2002 – Addison Wesley)

Artificial Intelligence: A New Synthesis, by N.J. Nilsson (1998 – Morgan Kauffman)

Artificial Intelligence: A Modern Approach, by Stuart Russell and Peter Norvig (1995 – Prentice Hall)

The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics, by Roger Penrose (1989 – Oxford University Press)

Understanding Intelligence, by Rolf Pfeiffer and Christian Scheier (2000 – The MIT Press)

Artificial Intelligence: An Engineering Approach, by Robert J. Schalkoff (1990 – McGraw Hill)

The Encyclopedia of Artificial Intelligence, edited by S.C. Shapiro (1992 - Wiley)

Artificial Intelligence, by Patrick Henry Winston (1992 – Addison Wesley)