Part I

Beginnings

1

Dreams and Dreamers

THE QUEST FOR ARTIFICIAL INTELLIGENCE (AI) BEGINS WITH DREAMS – AS ALL quests do. People have long imagined machines with human abilities – automata that move and devices that reason. Human-like machines are described in many stories and are pictured in sculptures, paintings, and drawings.

You may be familiar with many of these, but let me mention a few. The *Iliad of Homer* talks about self-propelled chairs called "tripods" and golden "attendants" constructed by Hephaistos, the lame blacksmith god, to help him get around.[1*] And, in the ancient Greek myth as retold by Ovid in his *Metamorphoses*, Pygmalian sculpts an ivory statue of a beautiful maiden, Galatea, which Venus brings to life:[2]

> The girl felt the kisses he gave, blushed, and, raising her bashful eyes to the light, saw both her lover and the sky.

The ancient Greek philosopher Aristotle (384–322 BCE) dreamed of automation also, but apparently he thought it an impossible fantasy – thus making slavery necessary if people were to enjoy leisure. In his *The Politics*, he wrote[3]

> For suppose that every tool we had could perform its task, either at our bidding or itself perceiving the need, and if – like . . . the tripods of Hephaestus, of which the poet [that is, Homer] says that "self-moved they enter the assembly of gods" – shuttles in a loom could fly to and fro and a plucker [the tool used to pluck the strings] play a lyre of their own accord, then master craftsmen would have no need of servants nor masters of slaves.

Aristotle might have been surprised to see a Jacquard loom weave of itself or a player piano doing its own playing.

Pursuing his own visionary dreams, Ramon Llull (circa 1235–1316), a Catalan mystic and poet, produced a set of paper discs called the *Ars Magna* (Great Art), which was intended, among other things, as a debating tool for winning Muslims to the Christian faith through logic and reason. (See Fig. 1.1.) One of his disc assemblies was inscribed with some of the attributes of God, namely goodness, greatness, eternity, power, wisdom, will, virtue, truth, and glory. Rotating the discs appropriately was supposed to produce answers to various theological questions.[4]

Ahead of his time with inventions (as usual), Leonardo Da Vinci sketched designs for a humanoid robot in the form of a medieval knight around the year 1495. (See Fig. 1.2.) No one knows whether Leonardo or contemporaries tried to build his

[*] So as not to distract the general reader unnecessarily, numbered notes containing citations to source materials appear at the end of each chapter. Each of these is followed by the number of the page where the reference to the note occurred.

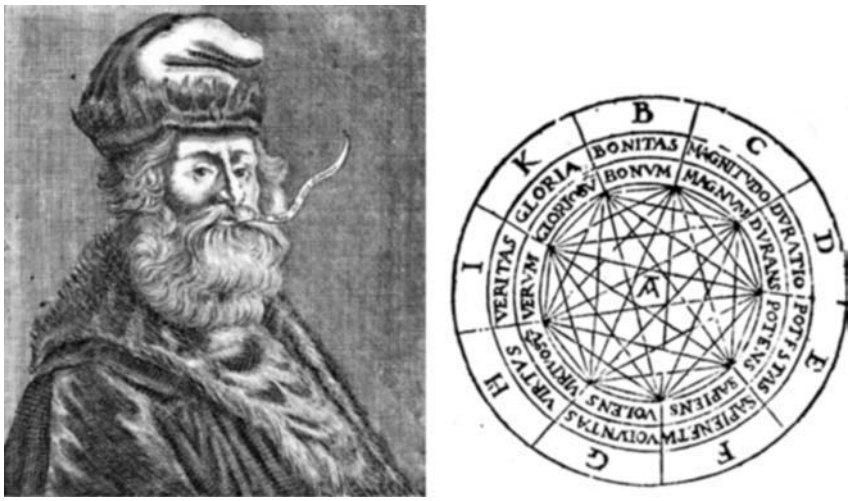

Figure 1.1. Ramon Llull (left) and his *Ars Magna* (right).

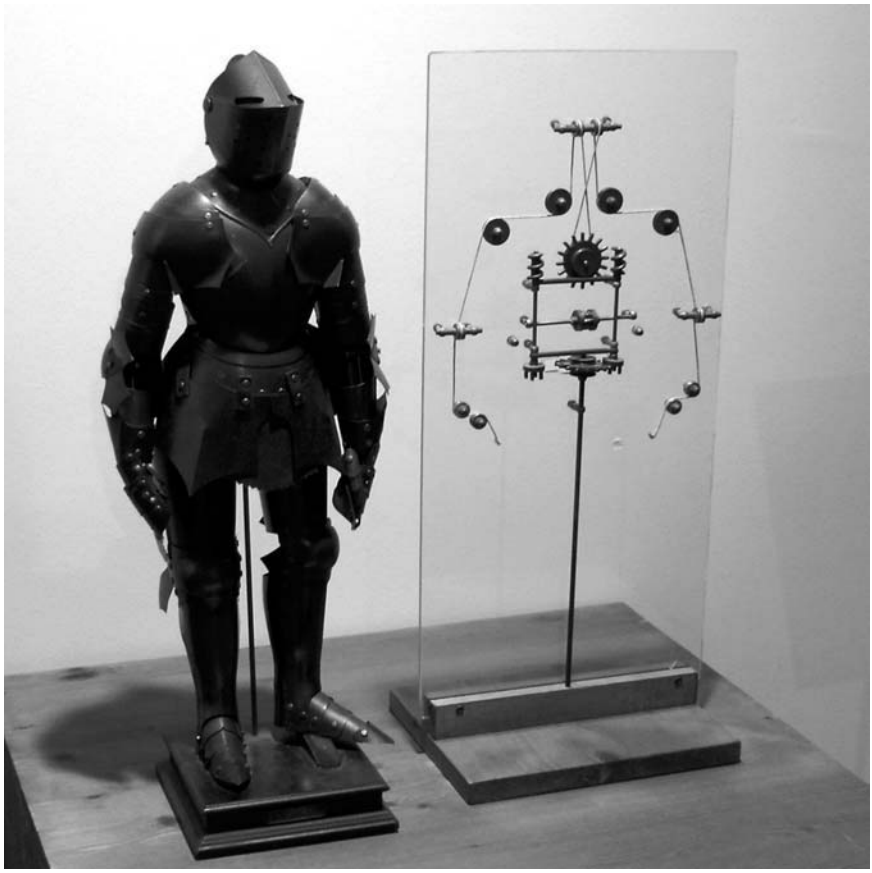

Figure 1.2. Model of a robot knight based on drawings by Leonardo da Vinci.

design. Leonardo's knight was supposed to be able to sit up, move its arms and head, and open its jaw.[5]

The Talmud talks about holy persons creating artificial creatures called "golems." These, like Adam, were usually created from earth. There are stories about rabbis using golems as servants. Like the Sorcerer's Apprentice, golems were sometimes difficult to control.

In 1651, Thomas Hobbes (1588–1679) published his book *Leviathan* about the social contract and the ideal state. In the introduction, Hobbes seems to say that it might be possible to build an "artificial animal."[6]

> For seeing life is but a motion of limbs, the beginning whereof is in some principal part within, why may we not say that all automata (engines that move themselves by springs and wheels as doth a watch) have an artificial life? For what is the heart, but a spring; and the nerves, but so many strings; and the joints, but so many wheels, giving motion to the whole body . . .

Perhaps for this reason, the science historian George Dyson refers to Hobbes as the "patriarch of artificial intelligence."[7]

In addition to fictional artifices, several people constructed actual automata that moved in startlingly lifelike ways.[8] The most sophisticated of these was the mechanical duck designed and built by the French inventor and engineer, Jacques de Vaucanson (1709–1782). In 1738, Vaucanson displayed his masterpiece, which could quack, flap its wings, paddle, drink water, and eat and "digest" grain.

As Vaucanson himself put it,[9]

> My second Machine, or *Automaton*, is a *Duck*, in which I represent the Mechanism of the Intestines which are employed in the Operations of Eating, Drinking, and Digestion: Wherein the Working of all the Parts necessary for those Actions is exactly imitated. The Duck stretches out its Neck to take Corn out of your Hand; it swallows it, digests it, and discharges it digested by the usual Passage.

There is controversy about whether or not the material "excreted" by the duck came from the corn it swallowed. One of the automates-anciens Web sites[10] claims that "In restoring Vaucanson's duck in 1844, the magician Robert-Houdin discovered that 'The discharge was prepared in advance: a sort of gruel composed of green-coloured bread crumb . . . '."

Leaving digestion aside, Vaucanson's duck was a remarkable piece of engineering. He was quite aware of that himself. He wrote[11]

> I believe that Persons of Skill and Attention, will see how difficult it has been to make so many different moving Parts in this small *Automaton*; as for Example, to make it rise upon its Legs, and throw its Neck to the Right and Left. They will find the different Changes of the *Fulchrum's* or Centers of Motion: they will also see that what sometimes is a Center of Motion for a moveable Part, another Time becomes moveable on that Part, which Part then becomes fix'd. In a Word, they will be sensible of a prodigious Number of Mechanical Combinations.
>
> This Machine, when once wound up, performs all its different Operations without being touch'd any more.
>
> I forgot to tell you, that the *Duck* drinks, plays in the Water with his Bill, and makes a gurgling Noise like a real living *Duck*. In short, I have endeavor'd to make it imitate all the Actions of the living Animal, which I have consider'd very attentively.

Figure 1.3. Frédéric Vidoni's ANAS, inspired by Vaucanson's duck. (Photograph courtesy of Frédéric Vidoni.)

Unfortunately, only copies of the duck exist. The original was burned in a museum in Nijninovgorod, Russia around 1879. You can watch, ANAS, a modern version, performing at http://www.automates-anciens.com/video_1/duck_automaton_vaucanson_500.wmv.[12] It is on exhibit in the Museum of Automatons in Grenoble and was designed and built in 1998 by Frédéric Vidoni, a creator in mechanical arts. (See Fig. 1.3.)

Figure 1.4. A scene from a New York production of *R.U.R.*

Returning now to fictional automata, I'll first mention the mechanical, life-sized doll, Olympia, which sings and dances in Act I of *Les Contes d'Hoffmann* (*The Tales of Hoffmann*) by Jacques Offenbach (1819–1880). In the opera, Hoffmann, a poet, falls in love with Olympia, only to be crestfallen (and embarrassed) when she is smashed to pieces by the disgruntled Coppélius.

A play called *R.U.R.* (*Rossum's Universal Robots*) was published by Karel Čapek (pronounced CHAH pek), a Czech author and playwright, in 1920. (See Fig. 1.4.) Čapek is credited with coining the word "robot," which in Czech means "forced labor" or "drudgery." (A "robotnik" is a peasant or serf.)

The play opened in Prague in January 1921. The Robots (always capitalized in the play) are mass-produced at the island factory of Rossum's Universal Robots using a chemical substitute for protoplasm. According to a Web site describing the play,[13] "Robots remember everything, and think of nothing new. According to Domin [the factory director] 'They'd make fine university professors.' . . . once in a while, a Robot will throw down his work and start gnashing his teeth. The human managers treat such an event as evidence of a product defect, but Helena [who wants to liberate the Robots] prefers to interpret it as a sign of the emerging soul."

I won't reveal the ending except to say that Čapek did not look eagerly on technology. He believed that work is an essential element of human life. Writing in a 1935 newspaper column (in the third person, which was his habit) he said: "With outright horror, he refuses any responsibility for the thought that machines could take the place of people, or that anything like life, love, or rebellion could ever awaken in their cogwheels. He would regard this somber vision as an unforgivable overvaluation of mechanics or as a severe insult to life."[14]

There is an interesting story, written by Čapek himself, about how he came to use the word robot in his play. While the idea for the play "was still warm he rushed immediately to his brother Josef, the painter, who was standing before an easel and painting away. . . . 'I don't know what to call these artificial workers,' he said. 'I could

call them Labori, but that strikes me as a bit bookish.' 'Then call them Robots,' the painter muttered, brush in mouth, and went on painting."[15]

The science fiction (and science fact) writer Isaac Asimov wrote many stories about robots. His first collection, *I, Robot,* consists of nine stories about "positronic" robots.[16] Because he was tired of science fiction stories in which robots (such as Frankenstein's creation) were destructive, Asimov's robots had "Three Laws of Robotics" hard-wired into their positronic brains. The three laws were the following:

FIRST LAW: A robot may not injure a human being, or, through inaction, allow a human being to come to harm.

SECOND LAW: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

THIRD LAW: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Asimov later added a "zeroth" law, designed to protect humanity's interest:[17]

ZEROTH LAW: A robot may not injure humanity, or, through inaction, allow humanity to come to harm.

The quest for artificial intelligence, quixotic or not, begins with dreams like these. But to turn dreams into reality requires usable clues about how to proceed. Fortunately, there were many such clues, as we shall see.

Notes

1. *The Iliad of Homer*, translated by Richmond Lattimore, p. 386, Chicago: The University of Chicago Press, 1951. (Paperback edition, 1961.) [3]
2. Ovid, *Metamorphoses*, Book X, pp. 243–297, from an English translation, circa 1850. See http://www.pygmalion.ws/stories/ovid2.htm. [3]
3. Aristotle, *The Politics*, p. 65, translated by T. A. Sinclair, London: Penguin Books, 1981. [3]
4. See E. Allison Peers, *Fool of Love: The Life of Ramon Lull*, London: S. C. M. Press, Ltd., 1946. [3]
5. See http://en.wikipedia.org/wiki/Leonardo's_robot. [5]
6. Thomas Hobbes, *The Leviathon*, paperback edition, Kessinger Publishing, 2004. [5]
7. George B. Dyson, *Darwin Among the Machines: The Evolution of Global Intelligence*, p. 7, Helix Books, 1997. [5]
8. For a Web site devoted to automata and music boxes, see http://www.automates-anciens.com/english_version/frames/english_frames.htm. [5]
9. From Jacques de Vaucanson, "An account of the mechanism of an automaton, or image playing on the German-flute: as it was presented in a memoire, to the gentlemen of the Royal-Academy of Sciences at Paris. By M. Vaucanson . . . Together with a description of an artificial duck. . . . " Translated out of the French original, by J. T. Desaguliers, London, 1742. Available at http://e3.uci.edu/clients/bjbecker/NatureandArtifice/week5d.html. [5]
10. http://www.automates-anciens.com/english_version/automatons-music-boxes/vaucanson-automatons-androids.php. [5]
11. de Vaucanson, Jacques, *op. cit.* [5]

12. I thank Prof. Barbara Becker of the University of California at Irvine for telling me about the automates-anciens.com Web sites. [6]
13. http://jerz.setonhill.edu/resources/RUR/index.html. [7]
14. For a translation of the column entitled "The Author of Robots Defends Himself," see http://www.depauw.edu/sfs/documents/capek68.htm. [7]
15. From one of a group of Web sites about Čapek, http://Capek.misto.cz/english/robot.html. See also http://Capek.misto.cz/english/. [8]
16. The Isaac Asimov Web site, http://www.asimovonline.com/, claims that "Asimov did not come up with the title, but rather his publisher 'appropriated' the title from a short story by Eando Binder that was published in 1939." [8]
17. See http://www.asimovonline.com/asimov_FAQ.html#series13 for information about the history of these four laws. [8]

2

Clues

CLUES ABOUT WHAT MIGHT BE NEEDED TO MAKE MACHINES INTELLIGENT ARE scattered abundantly throughout philosophy, logic, biology, psychology, statistics, and engineering. With gradually increasing intensity, people set about to exploit clues from these areas in their separate quests to automate some aspects of intelligence. I begin my story by describing some of these clues and how they inspired some of the first achievements in artificial intelligence.

2.1 From Philosophy and Logic

Although people had reasoned logically for millennia, it was the Greek philosopher Aristotle who first tried to analyze and codify the process. Aristotle identified a type of reasoning he called the *syllogism* ". . . in which, certain things being stated, something other than what is stated follows of necessity from their being so."[1]

Here is a famous example of one kind of syllogism:[2]

1. *All humans are mortal.* (stated)
2. *All Greeks are humans.* (stated)
3. *All Greeks are mortal.* (result)

The beauty (and importance for AI) of Aristotle's contribution has to do with the *form* of the syllogism. We aren't restricted to talking about humans, Greeks, or mortality. We could just as well be talking about something else – a result made obvious if we rewrite the syllogism using arbitrary symbols in the place of *humans*, *Greeks*, and *mortal*. Rewriting in this way would produce

1. *All B's are A.* (stated)
2. *All C's are B's.* (stated)
3. *All C's are A.* (result)

One can substitute anything one likes for A, B, and C. For example, *all athletes are healthy* and *all soccer players are athletes*, and therefore *all soccer players are healthy*, and so on. (Of course, the "result" won't necessarily be true unless the things "stated" are. Garbage in, garbage out!)

Aristotle's logic provides two clues to how one might automate reasoning. First, patterns of reasoning, such as syllogisms, can be economically represented as *forms* or *templates*. These use generic symbols, which can stand for many different concrete instances. Because they can stand for anything, the symbols themselves are unimportant.

Figure 2.1. Gottfried Leibniz.

Second, after the general symbols are replaced by ones pertaining to a specific problem, one only has to "turn the crank" to get an answer. The use of general symbols and similar kinds of crank-turning are at the heart of all modern AI reasoning programs.

In more modern times, Gottfried Wilhelm Leibniz (1646–1716; Fig. 2.1) was among the first to think about logical reasoning. Leibniz was a German philosopher, mathematician, and logician who, among other things, co-invented the calculus. (He had lots of arguments with Isaac Newton about that.) But more importantly for our story, he wanted to mechanize reasoning. Leibniz wrote[3]

> It is unworthy of excellent men to lose hours like slaves in the labor of calculation which could safely be regulated to anyone else if machines were used.

and

> For if praise is given to the men who have determined the number of regular solids . . . how much better will it be to bring under mathematical laws human reasoning, which is the most excellent and useful thing we have.

Leibniz conceived of and attempted to design a language in which all human knowledge could be formulated – even philosophical and metaphysical knowledge. He speculated that the propositions that constitute knowledge could be built from a smaller number of primitive ones – just as all words can be built from letters in an alphabetic language. His *lingua characteristica* or universal language would consist of these primitive propositions, which would comprise an *alphabet for human thoughts*.

The alphabet would serve as the basis for automatic reasoning. His idea was that if the items in the alphabet were represented by numbers, then a complex proposition could be obtained from its primitive constituents by multiplying the corresponding numbers together. Further arithmetic operations could then be used to determine whether or not the complex proposition was true or false. This whole process was to be accomplished by a *calculus ratiocinator* (calculus of reasoning). Then, when