CHAPTER 5

CYBERNETICS AND THE ORIGINS OF ARTIFICIAL INTELLIGENCE

1. Introduction. The turning point

In the previous chapters we identified various stages in the discovery of a novel strategy in the study of mind and behavior, suggested by 'new' notions of the machine. More than once we alluded to cybernetics, especially in Chapter 4. We mentioned there claims made by Rosenblueth, Wiener and Bigelow in their 1943 paper, emphasizing the role that began to be attributed to feedback in the study of various aspects of organism behavior. That paper is usually considered the manifesto of what Wiener called cybernetics, i.e. the study of "control and communication in the animal and the machine" (Wiener, 1948/1961).

But "Behavior, Purpose, and Teleology" was not an isolated herald of cybernetics. In 1943, Craik published *The Nature of Explanation*, and in the same year Warren McCulloch and Walter Pitts published a memorable article that would deeply influence both cybernetics and computer science (McCulloch and Pitts, 1943). "Certain ideas were in the air," was Seymour Papert's comment (Papert, 1968: 136). This is even clearer if one considers Ashby's 1940 article mentioned earlier and, in a different context, Jean Piaget's analysis of equilibration processes (Piaget, 1941), which broadly anticipated the concepts and methods of cybernetics, whose terminology Piaget was later to adopt.

As McCulloch would later recall, at that time he and Pitts were unaware of the results that Claude Shannon had already published (Shannon, 1938). And yet, McCulloch and Pitts applied the same tool used by Shannon, Boolean algebra, to investigate a quite different domain. They introduced networks consisting of "formal" neurons, simplified analogs of the neurons in the brain, functioning according to the all-or-nothing law (a neuron does or does not fire according to whether the intensity of the impulses it receives does or does not reach a certain threshold), whereas Shannon modeled the components of electric circuits, functioning according to a similar law (a relay closes or opens according to whether the current does or does not reach a certain intensity). Shannon's work would prove decisive in designing the circuits of digital computers. Though they knew nothing of Shannon's work, McCulloch and Pitts were perfectly aware of Alan Turing's. In 1936, Turing introduced the abstract computing machines that bear his name, and explicitly construed a universal machine which could simulate, with appropriate encoding, any computation carried out by any Turing machine (including, of course, the universal one) (Turing, 1936-37). These Turing machines came to be considered as the more general abstract model of every physically realized general-purpose digital computer. McCulloch and Pitts concluded that every net of formal neurons, if furnished with a tape and suitable input, output and scanning systems, is equivalent to a Turing machine (McCulloch and Pitts, 1943: 132).

From the early 1940s through to the mid-1950s, many other events signaled a turning point in the very understanding of the concept of machine, in the construction of actual machines and in man-machine interaction. Let us briefly consider some of those events.[1]

In 1941 Konrad Zuse built the Z3 electro-mechanical digital computer in Germany, and in 1944 Howard Aiken built another, the Mark I, in the United States. Research on wartime applications of computers was funded by generous as well as self-serving grants, and was repaid by rapid advances. While the defeat of Germany interrupted Zuse's work, in England and the United States the realization of large digital computers was relentlessly pursued, through the mobilization of extraordinary talents and resources. Already in 1943 the COLOSSUS computers, used to break German military codes, were functioning in England. Though these machines were dedicated to performing this kind of task, they were more evolved than the aforementioned computers—they were, among other things, completely electronic, i.e. with vacuum tubes replacing the electro-mechanical relays. This technological progress would make data processing really fast for the first time, leading to so-called first-generation computers. As they were blanketed by the strictest military secrecy, it was only from the mid-1970s that one began to learn about the characteristics of these machines, designed and built by a group headed by the mathematician Max Newman, and including I.J. Good and Donald Michie. Turing himself contributed to breaking the code of the legendary German machine ENIGMA. In the second half of the1940s, he would participate in two different projects for large computers: ACE (Automatic Computing Engine) at Teddington and MADM (Manchester Automatic Digital Machine) at Manchester.

In the United States, the construction of an electronic computer was completed in 1946. Its designers, J. Presper Eckert and John Mauchly, from the University of Pennsylvania, called it the Electronic Numerical Integrator and Calculator, or ENIAC. It was the largest computer ever built, and it is usually considered the first general-purpose computer. It was precisely within the ENIAC group that one of the most decisive events of these turning-point years came about. One of the consultants to the ENIAC project was John von Neumann. Few texts are as famous in the history of computer science as the *First Draft*, written by von Neumann in 1945. There, adopting McCulloch and Pitts' symbolism, he described the architecture of a newly-conceived general-purpose computer, which would remain basically unchanged over the years to come. Not only were data stored in its main memory; instructions to handle data, or programs, were stored there, too, and thus became as readily modifiable as the data themselves. The first large stored-program general-purpose computer, however, was built in Cambridge, England, by a group headed by Maurice Wilkes at the Mathematical Laboratory: in 1949 they completed the EDSAC (Electronic Delay Storage Automatic Calculator). In the United States, a stored-program general-purpose computer, the EDVAC (Electronic Discrete Variable Automatic Computer), was completed, on the design of von Neumann and others, in 1951.

Between 1954 and 1955, the first programming languages were introduced. After the Turing machine and von Neumann's architecture, they are, according to Allen Newell and

[1] For events, characters and texts in the history of computer science that are mentioned below, see Goldstine (1981), Hodges (1983), Metropolis et al. (1980), Randell (1973). See also the web site http://vlmp.museophile.com/computing.html.

Herbert Simon, the third crucial milestone in the history of computer science as it progressed towards AI (Newell e Simon, 1976). This new discipline, officially launched at the now celebrated Dartmouth seminar in 1956, aimed at reproducing by means of computer programs behaviors that, if observed in human beings, would be considered 'intelligent.'

The themes that will be dealt with in the main body of this chapter can be traced back to these turning-point years. First of all, we examine some artifacts of cybernetic robotics, attempting to identify their engineering significance, since they embody automatic control and electronic principles, and their theoretical significance, since they seem to embody some of the principles of the behavior of living organisms as well. This twofold source of interest, as we will see, remains a constant in the history of artificial modeling up to our day. Then, after isolating some stages in the ongoing influence of information-processing concepts over the psychology of the early 1950s, we will follow the coming of age of the 'intelligent' computer up to the first appearances of AI and of what Newell and Simon called Information Processing Psychology. Our chief goal, in this and the next chapter, is to emphasize, together with the radical novelties, the elements that these investigations share with certain experiments, methodological perspectives and philosophical assumptions whose discussion occupied us in the previous chapters.

2. Cybernetic robotics

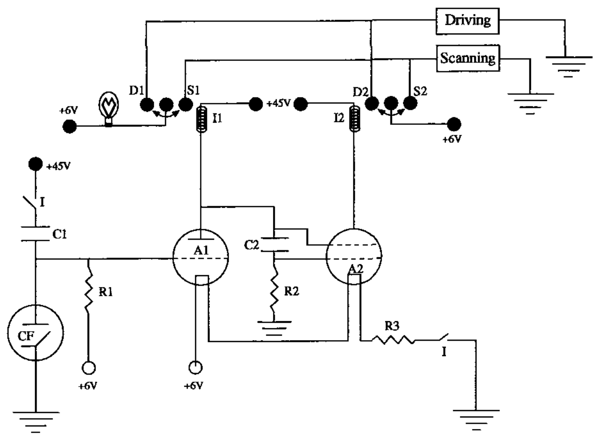
The technology underlying Hammond and Miessner's automaton, the phototropic dog Seleno whose performance had so struck Loeb, was exploited for wartime applications during the First World Word.[2] The circumstances surrounding some of the best-known cybernetic synthetic animals was no different. Notwithstanding the toy-like character of Grey Walter's *Machina speculatrix*, or "tortoise," one should not forget that two of its main characteristics, negative feedback and scanning, were at the center of experimentation that matured in the context of the Second World War. The first exemplars of the *Machina speculatrix* were implemented towards the end of the 1940s at Bristol, where Grey Walter worked in the Physiology Department of the Burden Neurological Institute. The tortoises were undoubtedly the most popular cybernetic robots, and were exhibited on several occasions (Walter, 1950; 1951).[3]

The *Machina speculatrix* shared with other robots of the period the appearance of a three-wheeled cart. The front wheel was both for propulsion and for steering. Furnished with a photoelectric cell (its first sensor) mounted on the cart, the machine manifested a form of positive tropism. In the dark it explored its surroundings by circular movements, but quickly headed towards any light source striking its photoelectric cell. The variable behavior of this machine depended on the combined operation of two motors (the circuit in Plate 5.1 accounts for such operation). Unlike Hammond and Miessner's phototropic

[2] See Chapter 1, pp. 5-6. There are reports of similar automata presented at Paris in 1929 and at New York in 1939 (de Latil, 1953: 266). Devices exemplified by these artifacts, mounted on torpedoes, were used by Germans against Allied ships during the Second World War (Nemes, *1969: 164*).

[3] An exemplar has been restored by Owen Holland (see the web page of the Intelligent Autonomous Systems Engineering Laboratory of Bristol at http://www.uwe.ac.uk./facults/eng/ias).

Plate 5.1



In this simplified circuit diagram of Grey Walter's *Machina speculatrix* there are two distinct electrical circuits. The upper circuit feeds the two motors D (DRIVING) and S (SCANNING), which are activated by the double-throw circuit breakers of electromagnetic relays R1 and R2. The lower circuit controls the opening and closing of the two relays. This second circuit consists of two amplifiers, A1 and A2, activating relays R1 and R2. When A1 is conducting, relay R1 is closed; when A2 is conducting, R2 is closed.

The conductivity of amplifier A1 depends on the voltage present on its control grid. If it is greater than zero, A1 conducts; if, instead, it is close to or less than zero, A1 is cut off. To the grid are connected the resistance R1 and the photoelectric cell CF. In the dark, the photoelectric cell CF does not conduct, and the grid is polarized through R1, resting at a positive voltage. In the presence of light, CF conducts, the grid voltage goes down towards zero and A1 is cut off.

The conductivity of A2, on the other hand, depends both on the conductivity of A1 and on the R2-C2 circuit. Let us suppose that A1 is conducting: in this case C2 and the screen grid of A2 have a low positive voltage (~10 V), while on the control grid the voltage is zero, since R2 is connected to ground. Therefore, A2 conducts. If, in this phase, A1 stops conducting because light impinges on CF, then the voltage on C2 rises to ~45 V, and A2 keeps conducting. If, afterwards, A1 resumes conducting, then the voltage on C2 returns to ~10 V, but since the capacitator was charged earlier at 45 V, it now applies a negative pulse to the control grid of A2 while discharging through R2. During this discharge, A2 will be cut off (the control voltage will slowly rise from −35 V to 0 V). When the discharge is finished, A2 will start conducting again.

Finally, depending on the position of the relays, the machine exhibits different behaviors. For example, when the machine finds itself in the dark, the amplifiers A1 and A2 conduct, with the effect of closing the contacts S1 and D2, and the machine *goes forward* and *steers*. If the machine encounters a light, A1 stops conducting and R1 closes on D1, causing the machine to go *forward* without steering because contact D2 is active. If, during this movement, the machine once again finds itself in the dark, A1 starts conducting again and S1 closes once more, while A2, for a period of time t=~R2C2, is cut off, thus closing S2, and hence the machine *steers* without progressing. After the discharge, A2 starts conducting again and, therefore, again switching from S2 to D2, so that the machine *goes forward* and *steers*.

machine, Grey Walter's moved away from light, if the latter was too intense or bright, and stopped at a certain distance from it, thereby avoiding the fate of the moth in the candle, as Grey Walter remarked in his popular book, *The Living Brain* (Walter, 1953: 84). Furthermore, if placed before two lights of the same intensity, the machine did not show any sign of 'neurosis,' to use the term which, in the days of the robot approach, Ellson had used to describe certain behaviors of his learning model (Plate 3.4). Thanks to its scanning device, the machine went first towards one and then towards the other light source. If it bumped against an obstacle, a mechanical switch (its second sensor) detected the jolt and inserted a condenser between the output and the input of its central amplifier. It thus became an oscillator, suppressing its phototropic and scanning behaviors and giving way to a backward-forward and left-right movement that in the end enabled the machine to dodge the obstacle with a reasonable rate of success.

Much like Ellson's model, the machine reserved some surprises for its designer, as it produced a piece of behavior that had not been explicitly foreseen. For example, the tortoises manifested "self-recognition" (in front of a mirror) and "mutual recognition" (when two exemplars met each other), and these abilities were discovered accidentally, because a pilot-light mounted on each machine turned on to indicate when the steering-servo was in operation. Recalling the wartime applications of phototropic feedback devices, Grey Walter noted that the machine could be transformed into a self-directing missile, with some simple modifications and by inserting a second photoelectric cell (p. 202). In other words, the tortoise would lose its interesting exploratory abilities, and behave like Hammond and Miessner's phototropic dog.

Grey Walter was profuse in his use of psychological terms when he described his machine as "spontaneous," "purposeful," "independent" and able to "recognize" itself and other individuals of the same species, as well as to exhibit a food-searching instinct (it automatically went towards its 'nest' to recharge its batteries when they were about to run out). He admitted that all this could be seen as nothing but a series of tricks; nonetheless, an outside observer, let us say a biologist, would have used just this mentalistic terminology had he witnessed this behavior in real animals. In fact, Grey Walter pressed on to emphasize the fact that behaviors so close to those of a living organism had been obtained with a *minimal* number of units corresponding to nerve cells. In his robots there were two miniature valves, two relays, two sensors, one light-sensitive and the other touch-sensitive, and two effectors. The latter were the two small electric motors which enabled the machine to go forward and to steer, respectively.

The machine's circuitry did not include a true form of memory, except for a delay of about one second in returning the central circuit from its oscillator configuration, which allowed it to escape obstacles, to that of the amplifier, with which its tropistic and exploratory behavior resumed. What really mattered was the presence of a feedback loop in which the environment was one of the two components, the other being the machine. As for its memory characteristics, Grey Walter observed, his "two-element synthetic animals" were not, therefore, comparable with Ross' 1935 robot rat, which was equipped with a "memory disk" (see Figure 3.4). He then designed a more refined electrical circuit, with which his synthetic animals would have been able to 'remember' and 'forget.'

http://www.springer.com/978-1-4020-0606-7

The Discovery of the Artificial
Behavior, Mind and Machines Before and Beyond
Cybernetics
Cordeschi, R.
2002, XX, 314 p., Hardcover
ISBN: 978-1-4020-0606-7