CHAPTER 8

Boolean Algebra

Aristotle deserves full credit as the founder of formal logic even though he restricted his attention almost entirely to the syllogism. Today, when the syllogism has become a trivial part of logic, it is hard to believe that for 2,000 years it was the principal topic of logical studies, and that as late as 1797 Immanuel Kant could write that logic was "a closed and completed body of doctrine."

"In syllogistic inference," Bertrand Russell once explained, "you are supposed to know already that all men are mortal and that Socrates is a man; hence you deduce what you never suspected before, that Socrates is mortal. This form of inference does actually occur, though very rarely." Russell goes on to say that the only instance he ever heard of was prompted by a comic issue of *Mind*, a British philosophical journal, that the editors concocted as a special Christmas number in 1901. A German philosopher, puzzled by the magazine's advertisements, eventually reasoned: Everything in this magazine is a joke, the advertisements are in this magazine, therefore the advertisements are jokes. "If you wish to become a logician," Russell wrote elsewhere, "there is one piece of sound advice which I cannot urge too strongly, and that is: Do *not* learn the traditional logic. In Aristotle's day it was a creditable effort, but so was the Ptolemaic astronomy."

The big turning point came in 1847 when George Boole (1815–1864), a modest, self-taught son of a poor English shoemaker *[see Figure 40]*, published *The Mathematical Analysis of Logic*. This and other papers led to his appointment (although he had no university degree) as professor of mathematics at Queens College (now University College) at Cork in Ireland, where he wrote his treatise *An Investigation of the Laws of Thought, on Which are Founded the Mathematical Theories of Logic and Probabilities* (London, 1854). The basic idea—substituting symbols for all the words used in formal logic—had occurred to others before, but Boole was the first to produce a workable system. By and large, neither philosophers nor mathematicians of his century showed much interest in this remarkable achievement. Perhaps that was one reason for Boole's tolerant attitude toward mathematical eccentrics. He wrote an article about a Cork crank named John Walsh *(Philosophical Magazine,* November 1851) that Augustus De Morgan, in his *Budget of Paradoxes,* calls "the best biography of a single hero of the kind that I know."

Boole died of pneumonia at the age of 49. His illness was attributed to a chill that followed a lecture he gave in wet clothes after having been caught in the rain. He was survived by his wife and five daughters. Norman Gridgeman, writing "In Praise of Boole" (see bibliography), gives some fascinating details about the six ladies. Boole's wife, Mary Everest, wrote popular books about her husband's views on mathematics and education. One book is titled *The Philosophy and Fun of Algebra.* The oldest daughter, Mary, married Charles Hinton, a mathematician and bigamist who wrote the first novel about flatland (see Chapter 12 of my *Unexpected Hanging),* as well as books on the fourth dimension.

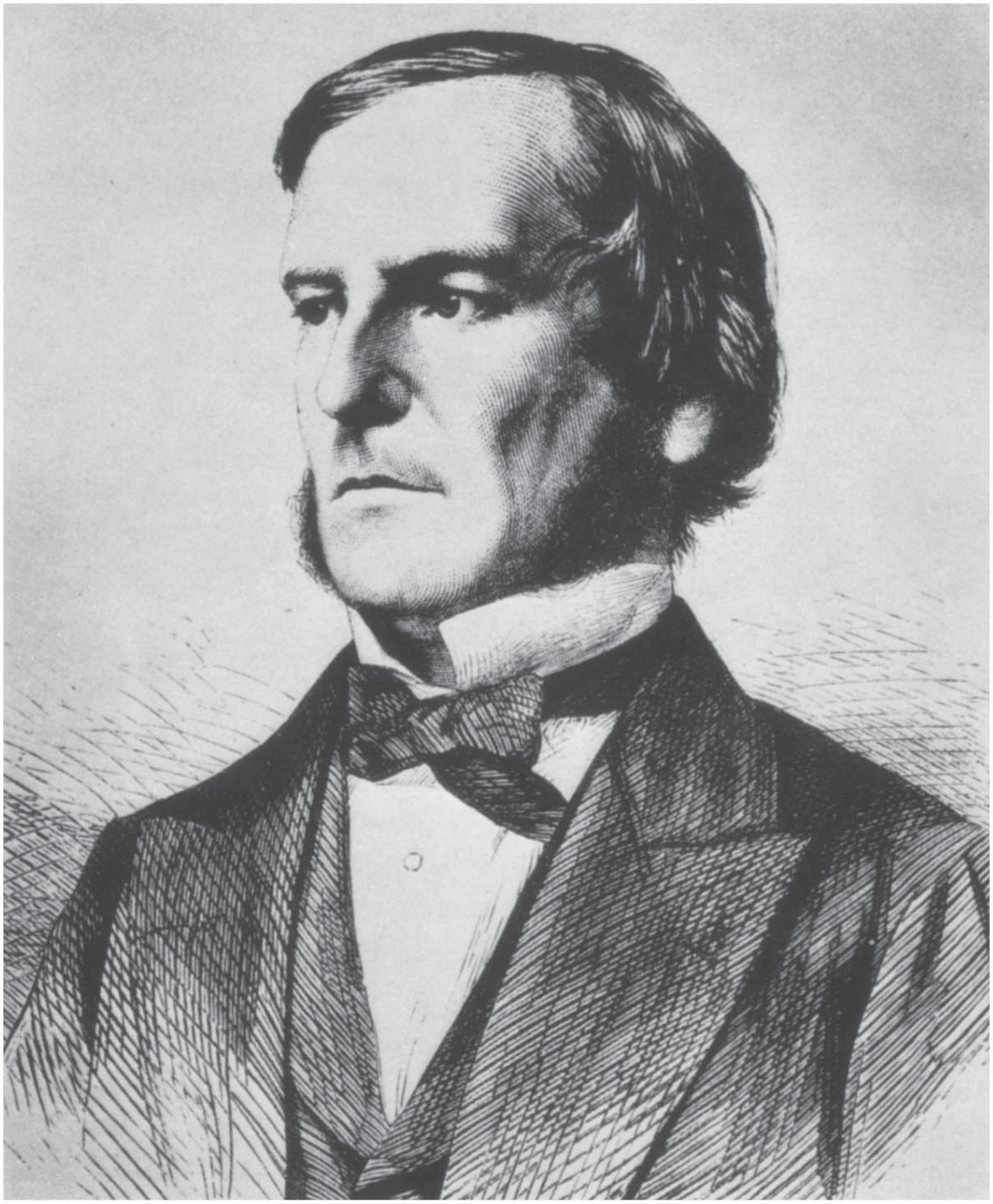


FIGURE 40
George Boole

Margaret became the mother of Sir Geoffrey Taylor, a Cambridge mathematician. Alicia, intrigued by Charles Hinton's excursions into higher space dimensions, made some significant discoveries in the field. Lucy became a professor of chemistry. Ethel Lilian, the youngest daughter, married Wilfrid Voynich, a Polish scientist. They settled in Manhattan, where Ethel died in 1960. She wrote several novels, including one called *The Gadfly* (1898) that became so popular in Russia that three operas were based on it. In recent years a million copies have been sold in China. "Modern Russians are constantly amazed," writes Gridgeman, "that so few Westerners have heard of E. L. Voynich, the great English novelist."

The few who appreciated Boole's genius (notably the German mathematician Ernst Schröder) rapidly improved on Boole's notation, which was clumsy mainly because of Boole's attempt to make his system resemble traditional algebra. Today Boolean algebra refers to an "uninterpreted" abstract structure that can be axiomized in all kinds of ways but that is essentially a streamlined, simplified version of Boole's system. "Uninterpreted" means that no meanings whatever—in logic, mathematics, or the physical world—are assigned to the structure's symbols.

As in the case of all purely abstract algebras, many different interpretations can be given to Boolean symbols. Boole himself interpreted his system in the Aristotelian way as an algebra of classes and their properties, but he greatly extended the old class logic beyond the syllogism's narrow confines. Since Boole's notation has been discarded, modern Boolean algebra is now written in the symbols of set theory, a set being the same as what Boole meant by a class: any collection of individual "elements." A set can be finite, such as the numbers 1, 2, 3, the residents of Omaha who have green eyes, the corners of a cube, the planets of the solar system, or any other specified collection of things. A set also can be infinite, such as the set of even integers or possibly the set of all stars. If we specify a set, finite or infinite, and then consider all its subsets (they include the set itself as well as the empty set of no members) as being related to one another by inclusion (that is, the set 1, 2, 3 is included in the set 1, 2, 3, 4, 5), we can construct a Boolean set algebra.

A modern notation for such an algebra uses letters for sets, subsets, or elements. The "universal set," the largest set being considered, is symbolized by **U**. The empty, or null, set is $\emptyset$. The "union" of sets a and b (everything in a and b) is symbolized by $\cup$, sometimes called a cup. (The union of 1, 2 and 3, 4, 5 is 1, 2, 3, 4, 5.) The "intersection" of sets a and b (everything common to a and b) is symbolized by $\cap$, sometimes called a cap. (The intersection of 1, 2, 3 and 3, 4, 5 is 3.) If two sets are identical (for example, the set of odd numbers is the same as the set of all integers with a remainder of 1 when divided by 2), this is symbolized by $=$. The "complement" of set a—all elements of the universal set that are not in a—is indicated by a'. (The complement of 1, 2, with respect to the universal set 1, 2, 3, 4, 5, is 3, 4, 5.) Finally, the basic binary relation of set inclusion is symbolized by ϵ; $a \epsilon b$ means that a is a member of b.

As a matter of historical interest, Boole's symbols included letters for elements, classes, and subclasses: 1 for the universal class; 0 for the null class; $+$ for class union (which he took in an "exclusive" sense to mean those elements of two classes that are *not* held in common; the switch to the "inclusive" sense, first made by the British logician and economist William Stanley Jevons, had so many advantages that later logicians adopted it); $\times$ for class intersection; $=$ for identity; and the minus sign, $-$, for the removal of one set from another. To show the complement of x, Boole wrote $1 - x$. He had no symbol for class inclusion but could express it in various ways such as $a \times b = a$, meaning that the intersection of a and b is identical with all of a.

The Boolean algebra of sets can be elegantly diagrammed with Venn circles (after the English logician John Venn), which are now being introduced in many elementary school classes. Venn circles are diagrams of an interpretation of Boolean algebra in the point-set topology of the plane. Let two overlapping circles symbolize the union of two sets [*see Figure 41*], which we here take to be the set of the 10 digits and the set of the first 10 primes. The area outside both circles is the universal set. It is usually enclosed in a rectangle shaded to indicate that it is the null set; it is empty because we are con-

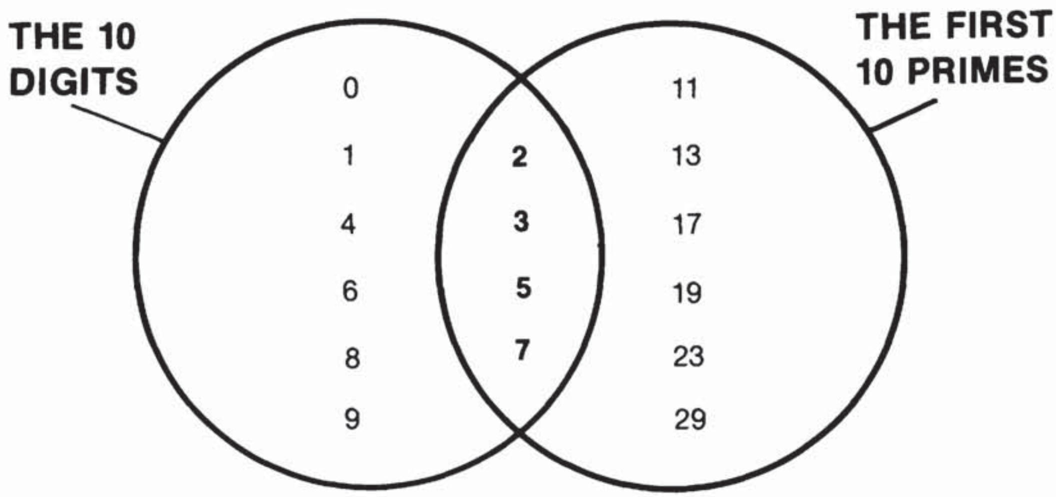
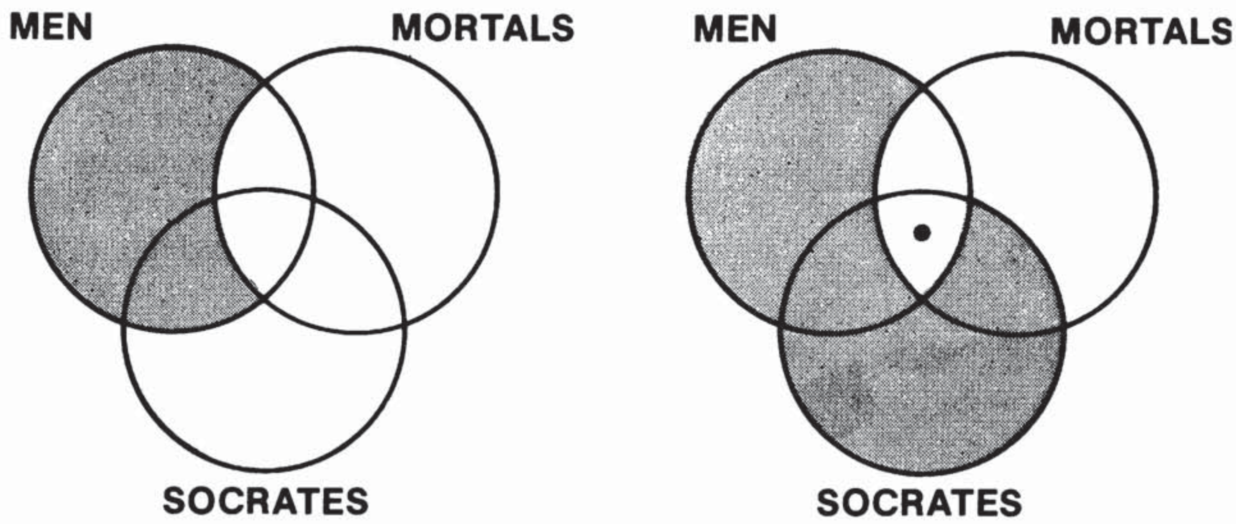


FIGURE 41
Venn diagram for set intersection

cerned solely with the elements inside the two circles. These 16 elements are the union of the two sets. The overlapping area contains the intersection. It consists of the set 2, 3, 5, 7: digits that are also among the first 10 primes.

Adopting the convention of shading any area known to represent an empty set, we can see how a three-circle Venn diagram proves the ancient syllogism Russell so scornfully cited. The circles are labeled to indicate sets of men, mortal things, and Socrates (a set with only one member). The first premise, "All men are mortal," is diagrammed by shading the men circle to show that the class of nonmortal men is empty [*see Figure 42, left*]. The second premise, "Socrates is a man," is similarly diagrammed by shading the Socrates circle to show that all of Socrates, namely himself, is inside the men circle [*see Figure 42, right*]. Now we inspect the diagram to see if the conclusion, "Socrates is mortal," is valid. It is. All of Socrates (the unshaded part of his circle marked by a dot) is inside the circle



Premise: "All men are mortal." *Premise: "Socrates is a man."*

FIGURE 42

BOOLEAN SET ALGEBRA	PROPOSITIONAL CALCULUS
U (UNIVERSAL SET)	T (TRUE)
ϕ (NULL SET)	F (FALSE)
$a, b, c, \ldots$ (SETS, SUBSETS, ELEMENTS)	$p, q, r, \ldots$ (PROPOSITIONS)
$a \cup b$ (UNION: ALL OF a AND b)	$p \vee q$ (DISJUNCTION: EITHER p ALONE OR q ALONE, OR BOTH, ARE TRUE.)
$a \cap b$ (INTERSECTION: WHAT a AND b HAVE IN COMMON)	$p \bullet q$ (CONJUNCTION: BOTH p AND q ARE TRUE.)
$a = b$ (IDENTITY: a AND b ARE THE SAME SET.)	$p \equiv q$ (EQUIVALENCE: IF AND ONLY IF p IS TRUE, THEN q IS TRUE.)
a' (COMPLEMENT: ALL OF **U** THAT IS NOT a)	$\sim p$ (NEGATION: p IS FALSE.)
$a \epsilon b$ (INCLUSION: a IS A MEMBER OF b.)	$p \supset q$ (IMPLICATION: IF p IS TRUE, q IS TRUE.)

FIGURE 43

Corresponding symbols in two versions of Boolean algebra of mortal things. By exploiting the topological properties of simple closed curves, we have a method of diagramming that is isomorphic with Boolean set algebra.

The first important new interpretation of Boolean algebra was suggested by Boole himself. He pointed out that if his 1 were taken as truth and his 0 as falsehood, the calculus could be applied to statements that are either true or false. Boole did not carry out this program but his successors did. It is now called the propositional calculus. This is the calculus concerned with true or false statements connected by such binary relations as "If p then q," "Either p or q but not both," "Either p or q or both," "If and only if p then q," "Not both p and q," and so on. The chart in Figure 43 shows the symbols of the propositional calculus that correspond to symbols for the Boolean set algebra.

It is easy to understand the isomorphism of the two interpretations by considering the syllogism about Socrates. Instead of saying, "All men are mortal," which puts it in terms of class properties or set inclusion, we rephrase it as, "If x is a man then x is a mortal." Now we are stating two propositions and joining them by the "connective" called "implication." This is diagrammed on Venn circles in exactly the same way we diagrammed "All men are mortal." Indeed, all the binary relations in the propositional calculus can be diagrammed with Venn circles and the circles can be used for solving simple problems in the calculus. It is shameful that writers of most introductory textbooks on formal logic have not yet caught on to this. They continue to use Venn circles to illustrate the old class-inclusion logic but fail to apply them to the propositional calculus, where they are just as efficient. Indeed, they are even more efficient, since in the propositional calculus one is unconcerned with the "existential quantifier," which asserts that a class is not empty because it has at least one member. This was expressed in the traditional logic by the word "some" (as in "Some apples are green"). To take care of such statements Boole had to tie his algebra into all sorts of complicated knots.

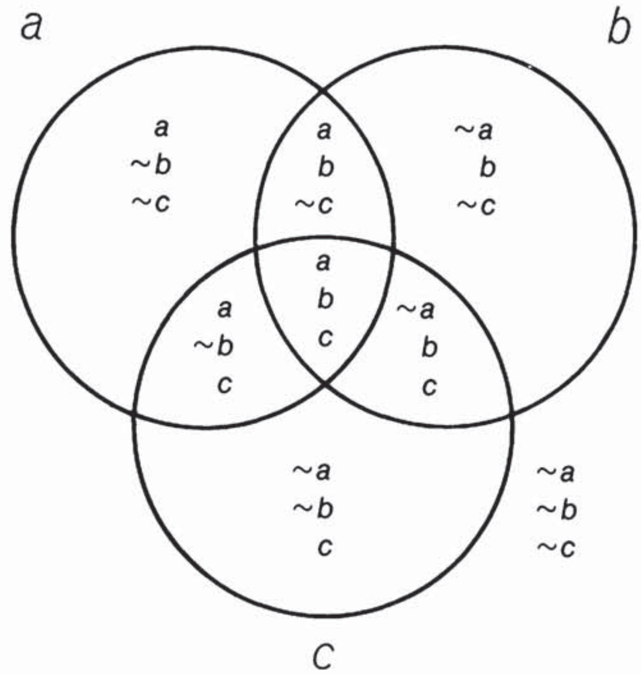


FIGURE 44
Venn diagram for martini puzzle

To see how easily the Venn circles solve certain types of logic puzzles, consider the following premises about three businessmen, Abner, Bill, and Charley, who lunch together every working day:

1. If Abner orders a martini, so does Bill.
2. Either Bill or Charley always orders a martini, but never both at the same lunch.
3. Either Abner or Charley or both always order a martini.
4. If Charley orders a martini, so does Abner.

To diagram these statements with Venn circles, we identify having a martini with truth and not having one with falsehood. The eight areas of the overlapping circles shown in Figure 44 are labeled to show all possible combinations of truth values for a, b, c, which stand for Abner, Bill, and Charley. Thus the area marked a, $\sim b$, c represents Abner's and Charley's having martinis while Bill does not. See if you can shade the areas declared empty by the four premises and then examine the result to determine who will order martinis if you lunch with the three men.

There are many other ways to interpret Boolean algebra. It can be taken as a special case of an abstract structure called a ring, or as a special case of another type of abstract structure called a lattice. It can be interpreted in combinatorial theory, information theory, graph theory, matrix theory, and metamathematical theories of deductive systems in general. In recent years the most useful interpretation has been in switching theory, which is important in the design of electronic computers but is not limited to electrical networks. It applies to any kind of energy transmission along channels with connecting devices that turn the energy on and off, or switch it from one channel to another.

The energy can be a flowing gas or liquid, as in modern fluid control systems [see "Fluid Control Devices," by Stanley W. Angrist, in *Scientific American*, December 1964]. It can be light beams. It can be mechanical energy as in the logic machine Jevons invented for solving four-term problems in Boolean algebra. It can be rolling marbles, as in several computerlike toys now on the market: Dr. Nim, Think-a-Dot, and Digi-Comp II. And if inhabitants of another planet have a highly developed sense of smell, their computers could use odors transmitted through tubes to sniffing outlets. As long as the energy either moves or does not move along a channel, there is an isomorphism between the two states and the two truth values of the propositional calculus. For every binary connective in the calculus, there is a corresponding switching circuit. Three simple examples are shown in Figure 45. The bottom circuit is used whenever two widely separated electric light switches are used to control one light. It is easy to see that if the light is off, changing the state of either switch will turn it on, and if the light is on, either switch will turn it off.

This electrical-circuit interpretation of Boolean algebra had been suggested in a Russian journal by Paul S. Ehrenfest as

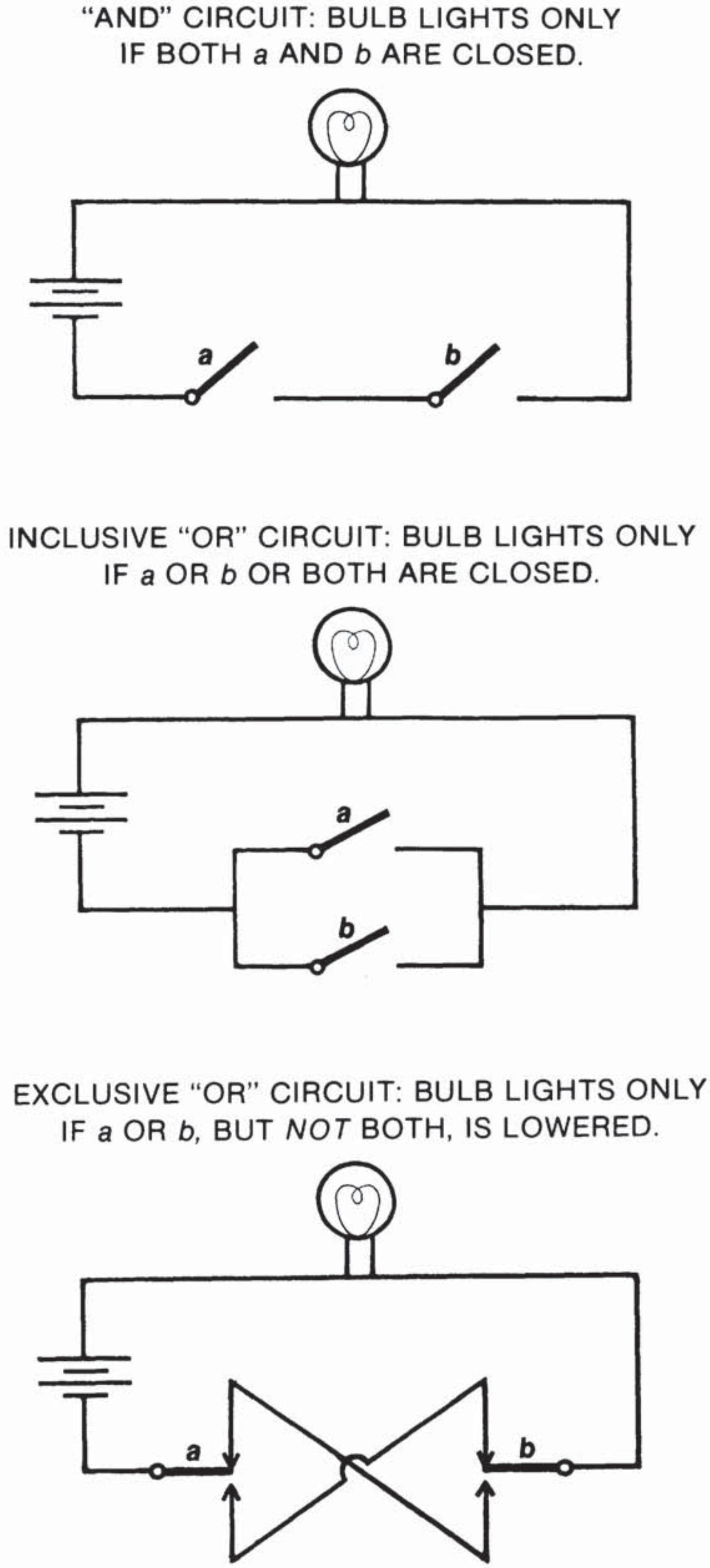


FIGURE 45

Circuits for three binary relations

early as 1910 and independently in Japan in 1936, but the first major paper, the one that introduced the interpretation to computer designers, was Claude E. Shannon's "A Symbolic Analysis of Relay and Switching Circuits" in the *Transactions of the American Institute of Electrical Engineers*, Vol. 57, December 1938. It was based on Shannon's 1937 master's thesis at the Massachusetts Institute of Technology.

Since Shannon's paper was published, Boolean algebra has become essential to computer design. It is particularly valuable in simplifying circuits to save hardware. A circuit is first translated into a statement in symbolic logic, the statement is "minimized" by clever methods, and the simpler statement is translated back to the design of a simpler circuit. Of course, in modern computers the switches are no longer magnetic devices or vacuum-tube diodes but transistors and other tiny semiconductors.

Now for one final interpretation of Boolean algebra that is a genuine curiosity. Consider the following set of eight numbers: 1, 2, 3, 5, 6, 10, 15, 30. They are the factors of 30, including 1 and 30 as factors. We interpret "union" as the least common multiple of any pair of those numbers. "Intersection" of a pair is taken to be their greatest common divisor. Set inclusion becomes the relation "is a factor of." The universal set is 30, the null set 1. The complement of a number a is $30/a$. With these novel interpretations of the Boolean relations, it turns out that we have a consistent Boolean structure! All the theorems of Boolean algebra have their counterparts in this curious system based on the factors of 30. For example, in Boolean algebra the complement of the complement of a is simply a, or in the propositional-calculus interpretation the negation of a negation is the same as no negation. More generally, only an odd series of negations equals a negation. Let us apply this Boolean law to the number 3. Its complement is $30/3 = 10$. The complement of 10 is $30/10 = 3$, which brings us back to 3 again.

Consider two famous Boolean laws called De Morgan's laws. In the algebra of sets they are

$$(a \cup b)' = a' \cap b'$$
$$(a \cap b)' = a' \cup b'.$$

In the propositional calculus they look like this:

$$\sim(a \vee b) \equiv \sim a \cdot \sim b$$
$$\sim(a \cdot b) \equiv \sim a \vee \sim b.$$

If the reader will substitute any two factors of 30 for a and b, and interpret the symbols as explained, he will find that De Morgan's laws hold. The fact that De Morgan's laws form a pair illustrates the famous duality principle of Boolean algebra. If in any statement you interchange union and intersection (if and wherever they appear) and interchange the universal and the null sets, and also reverse the direction of set inclusion, the result is another valid law. Moreover, these changes can be made all along the steps of the proof of one law to provide a valid proof of the other! (An equally beautiful duality principle holds in projective geometry with respect to interchanges of lines and points.)

The numbers 1, 2, 3, 5, 6, 7, 10, 14, 15, 21, 30, 35, 42, 70, 105, 210—the 16 factors of 210—also form a Boolean algebra when interpreted in the same way, although of course 210 is now the universal set and the complement of a is $210/a$. Can the reader discover a simple way to generate sets of 2^n numbers, where n is any positive integer, that will form Boolean systems of this peculiar kind?

ANSWERS

THREE VENN CIRCLES are shaded as in Figure 46 to solve the problem about the three men who lunch together. Each of the first four diagrams is shaded to represent one of the four premises of the problem. Superimposing the four to form the last diagram shows that if the four premises are true, the only possible

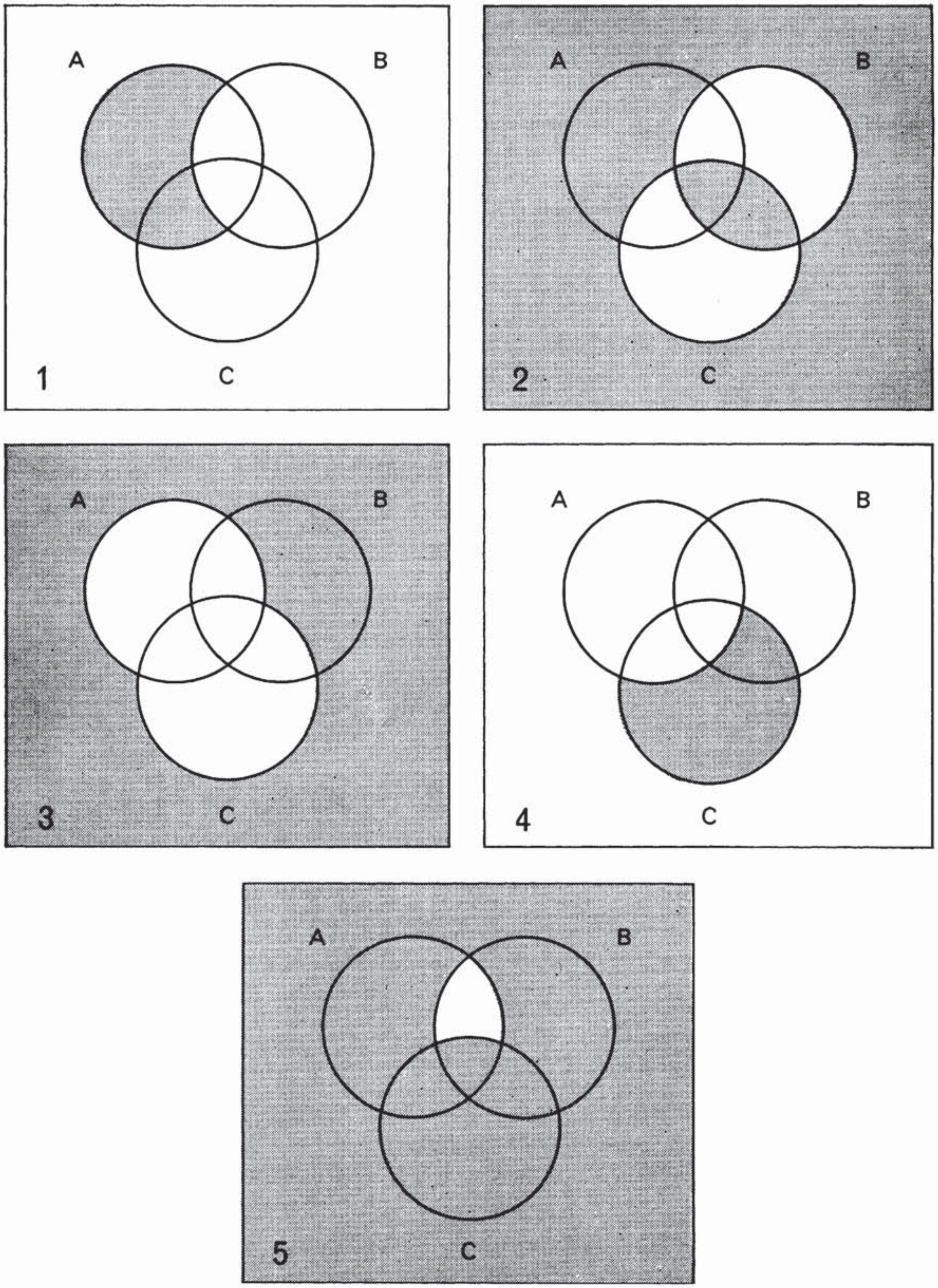


FIGURE 46
Venn-diagram solution to martini problem

combination of truth values is a, b, $\sim c$, or true a, true b, and false c. Since we are identifying truth with ordering a martini, this means that Abner and Bill always order martinis, whereas Charley never does.

The method of generating 2^n integers to form Boolean algebras was given by Francis D. Parker in *The American Mathematical Monthly* for March 1960, page 268. Consider a set of any number of distinct primes, say 2, 3, 5. Write down the multiples of all the subsets of these three primes, which include 0 (the null set) and the original set of three primes. Change 0 to 1. This produces the set 1, 2, 3, 5, 6, 10, 15, 30, the first of the examples given. In a similar way the four primes 2, 3, 5, 7 will generate the second example, the $2^4 = 16$ factors of 210. A proof that all such sets provide Boolean algebras can be found in *Boolean Algebra,* by R. L. Goodstein in the answer to problem No. 10, page 126.

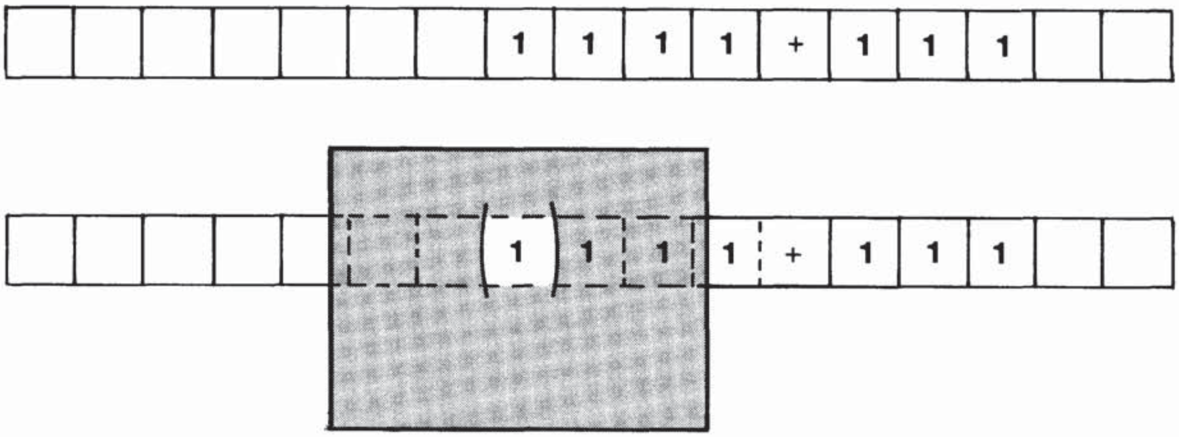
CHAPTER 9

Can Machines Think?

> *There was a time when it must have seemed highly improbable that machines should learn to make their wants known by sound, even through the ears of man; may we not conceive, then, that a day will come when those ears will be no longer needed, and the hearing will be done by the delicacy of the machine's own construction?—when its language shall have been developed from the cry of animals to a speech as intricate as our own?*
>
> —SAMUEL BUTLER, *Erewhon*

ALAN MATHISON TURING, a British mathematician who died in 1954 at the age of 42, was one of the most creative of the early computer scientists. Today he is best known for his concept of the Turing machine. We shall take a quick look at such machines and then consider one of Turing's less well-known ideas, the Turing game— a game that leads to deep and unsettled philosophical controversies.

A Turing machine is a "black box" (a machine with unspecified mechanisms) capable of scanning an infinite tape of square cells. The box can have any finite number of states. A finite portion of the tape consists of nonblank cells, each bearing one of a finite number of symbols. When the box views a cell, it can leave a symbol unaltered, erase it, erase it and print another symbol, or print a symbol in a blank cell. The tape is then shifted one cell to the left or right or stays fixed; the box either remains in the same state or clicks to a different state.



	STATE A	STATE B
1	1. ERASE THE 1. 2. SCAN NEXT CELL ON RIGHT. 3. GO TO STATE B.	1. SCAN NEXT CELL ON RIGHT. 2. STAY IN STATE B.
+		1. ERASE THE +. 2. PRINT 1. 3. STOP.

FIGURE 47
A Turing machine for addition

A table of rules describes what the box does for every achievable combination of symbol and state. Such a table completely defines a particular Turing machine. There is a countable (aleph null) infinity of Turing machines, each designed for a specific task, and for every task the machine's structure can vary widely in symbols, states, and rules.

A good way to grasp the essence of a Turing machine is to make one, albeit a trivial one [*see Figure 47*]. Eight cells on the paper tape are marked 1111 + 111, signifying the addition of 4 and 3 in the "unary" system in which an integer n is symbolized by n 1's. To make the machine, draw a small square (the black box) and cut two slits in it so that the tape can be inserted as shown. Adjust the tape so that the first 1 is visible. The table at the bottom of the picture gives all the necessary rules.

Start by assuming that the machine is in state A. Consult the table for the combination of symbol 1 and state A and do what it says: erase the 1, move the tape left (so that the box scans the next cell to the right) and assume that the machine clicks to state B. Continue in this way until the table tells you to stop.

If you follow the rules correctly, the machine will erase the first 1, shift the tape to the left cell by cell until it reaches the plus sign, change + to 1 and stop. The strip will then show 1111111, or 7. These simple rules obviously program the device to add any pair of positive integers, however large.

It is a tedious way to add, of course, but Turing's idea was to reduce machine computation to a simple and abstract schema, making it easier to analyze all kinds of thorny theoretical problems, such as what can and what cannot be computed. Turing showed that his idealized device can be programmed to do, in its clumsy way, anything the most powerful electronic computer can do. Like any computer—and like the human brain—it is limited by the fact that certain calculations (such as calculating pi) require an infinite number of steps and by the fact that some problems are unsolvable in principle; there *is* no algorithm, or effective procedure, by which they can be solved. A "universal Turing machine" is capable of doing whatever any special-purpose Turing machine can do. In brief, it computes anything that is computable.

In 1950 Turing's article "Computing Machinery and Intelligence" appeared in *Mind*, a British philosophical journal, and it has since been reprinted in several anthologies, including James R. Newman's *The World of Mathematics*. "I propose," Turing began, "to consider the question, 'Can machines think?' " This, Turing decided, was much too vague to have a meaningful answer. He proposed instead a related but more precise question: Can a computer be taught to win the "imitation game," now commonly called the Turing game or Turing test?

Turing based his test on a parlor game in which a man is concealed in one room and a woman in another. An interrogator of either sex asks the concealed players questions, which are conveyed by an intermediary; the answers are returned in typescript. Each player tries to convince the interrogator that he or she is, say, the woman. The interrogator wins if he guesses correctly who is telling the truth.

Suppose, Turing said, we replace one player with a learning machine that has been taught to converse in an ordinary language such as English. Is it possible for such a machine to deceive an interrogator when both the machine and its human partner try to persuade the questioner that he, she, or it is the human?

Several continuums blur the meaning of "deceive." How long a conversation is allowed? How intelligent is the interrogator? How intelligent is the person competing against the machine? It is possible today for a computer to pass the Turing test if the questioner is a child and is allowed only a few questions. There may be no dramatic future breakthroughs, just as there probably were no dramatic breakthroughs in the evolution of man. Conversation machines could improve gradually, so that it takes increasingly long dialogues, with increasingly clever interrogators, to beat the machines. Perhaps someday only an interrogating computer will consistently guess correctly. Turing himself made a cautious prediction. By the year 2000, he wrote, computers will speak fluently enough to deceive an "average interrogator" at least 30 percent of the time after "about five minutes" of dialogue. He imagined the following conversation as being typical:

Q. Please write me a sonnet on the subject of the Forth Bridge.

A. Count me out on this one. I never could write poetry.

Q. Add 34957 to 70764.

A. (Pause about 30 seconds and then give as answer) 105621.

Q. Do you play chess?

A. Yes.

Q. I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?

A. (After a pause of 15 seconds) R–R8 mate.

Many programs enabling computers to simulate English conversation now exist. Joseph Weizenbaum, an electrical engineer at the Massachusetts Institute of Technology, has written one he calls Eliza, after Eliza Doolittle of Shaw's *Pygmalion. The New York Times* (June 15, 1968) reported a typewritten dialogue with a woman in which an IBM 7094 computer assumed the role of a psychiatrist:

W. Men are all alike.

C. In what way?

W. They're always bugging us about something or other.

C. Can you think of a specific example?

W. Well, my boyfriend made me come here.

C. Your boyfriend made you come here?

W. He says I'm depressed much of the time.

C. I am sorry to hear you are depressed.

W. It's true. I am unhappy.

C. Do you think coming here will help you not to be unhappy?

W. I need some help, that much seems certain.

C. What would it mean to you if you got some help?

W. Perhaps I could learn to get along with my mother.

C. Tell me more about your family.

W. My mother takes care of me.

C. Who else in your family takes care of you?

W. My father.

C. Your father?

W. You are like my father in some ways.

C. What resemblance do you see?

W. You are not very aggressive but I think you don't want me to notice that.

C. What makes you think I am not very aggressive?
W. You don't argue with me.

The dialogue continues in a similar vein and is indistinguishable from a conversation between a patient and a nondirective therapist. The program was not even a learning program. Weizenbaum readily admitted that the computer did not "understand" anything being said. It certainly could not pass the Turing test. Suppose, however, that by the year 2000 computers can play the Turing game as well as they now play checkers and chess. What, if anything, will this reveal about the nature of the machine's "mind"?

Readers of Arthur C. Clarke's novel *2001: A Space Odyssey* may recall that HAL, the spaceship's talking computer, is said to "think" because he could "pass the Turing test with ease." (HAL stands for *h*euristically programmed *al*gorithmic computer, but Clarke may have had some trickier wordplay in mind when he picked the name. Can the reader figure out what it is?) Does HAL really think or does he just mimic thinking? Turing believed that when the time comes that computers converse well enough to pass his test, no one will hesitate to say that they are thinking.

Enormously tangled questions immediately arise. Can such a computer be self-conscious? Can it have emotions? A sense of humor? In short, should it be called a "person" or just a dead machine built to imitate a person? L. Frank Baum described Tiktok, a windup mechanical man, as a robot that "thinks, speaks, acts, and does everything but live."

Surely the ability of a computer to pass Turing tests would prove only that a computer could imitate human speech well enough to pass such tests. Suppose someone in the Middle Ages had proposed the following "tulip test." Will it ever be possible to make an artificial tulip that cannot be distinguished from a real tulip if one is allowed only to look at it? Fake tulips can now pass this test, but this tells us nothing about a chemist's ability to synthesize organic compounds or to make a tulip that will grow like a tulip. Just as we can touch what we think is a flower and exclaim, "Oh—it's artificial!" it seems unsurprising that a day might come when we can hold a lengthy conversation with what we think is a person, then open a door and be amazed to discover that we have been talking to a computer.

Keith Gunderson, in an important 1964 paper in which he criticized Turing for making too much of the significance of his test, expressed the point this way. "In the end, the steam drill outlasted John Henry as a digger of railway tunnels, but that didn't prove the machine had muscles; it proved that muscles were not necessary for digging railway tunnels."

A curious twist was given to the Turing test in a lecture by Michael Scriven, reprinted as "The Compleat Robot: A Prolegomena to Androidology" in *Dimensions of Mind*, edited by Sidney Hook. Scriven conceded that conversational ability does not prove that a computer possesses other attributes of a "person." Suppose, however, a conversing computer is taught the meaning of "truth" (in, say, the correspondence sense made precise by Alfred Tarski) and then is programmed so that it cannot lie. "This makes the robot unsuitable," Scriven said, "for use as a personal servant, advertising copywriter, or politician, but renders it capable of another service." We can now ask if it is aware that it exists, has emotions, thinks some jokes are funny, acts on its own free will, enjoys Keats and so on, and expect it to give correct answers.

There is the possibility that a "Scriven machine" (as it is called by one of several philosophers who criticize Scriven's paper in other chapters of Hook's anthology) will say no to all such questions. But if it gives yes answers, Scriven contends, we have as much justification for believing it as we have for believing a human, and no reason for not calling it a "person."

Philosophers disagree about Turing's and Scriven's arguments. In a short piece titled "The Supercomputer as Liar," Scriven replied to some of his critics. Mortimer J. Adler, in his book *The Difference of Man and the Difference It Makes*, takes the view that the Turing test is an "all-or-none affair," and that success or continued failure in creating computers capable of passing it will respectively weaken or strengthen the view that a man is radically different in kind from any possible machine as well as any subhuman animal.

Would conversing machines really alter the beliefs of people who hold such a view? It is not hard to imagine a television show 50 years from now in which guests ad-lib with a robot Johnny Carson whose memory has been stocked with a million jokes and that has been taught the art of timing by human comics. I doubt that anyone would suppose the computer had a sense of humor any more than a person defeated by a robot chess player supposes he has played against a machine radically different in kind from a computer that plays ticktacktoe. Rules of syntax and semantics are just not all that different from rules of chess.

At any rate, the debate continues, complicated by metaphysical and religious commitments and complex linguistic problems. All the age-old enigmas about mind and body and the nature of personality are being reformulated in a new terminology. It is hard to predict what thresholds will be crossed and how the crossings will affect fundamental philosophical disagreements as robots of the future improve—as they surely will—in their ability to think, speak, and act like humans.

Samuel Butler's chapters in *Erewhon* explaining why the Erewhonians destroyed their machines before the machines could become masters instead of servants were read 100 years ago as far-fetched satire. Today they read like sober prophecy. "There is no security," Butler wrote, "against the ultimate development of mechanical consciousness, in the fact of machines possessing little consciousness now. A mollusc has not much consciousness. Reflect upon the extraordinary advance which machines have made during the last few hundred years, and note how slowly the animal and vegetable kingdoms are advancing. The more highly organized machines are creatures not so much of yesterday, as of the last five minutes, so to speak, in comparison with past time."

ANSWERS

IF EACH LETTER in HAL is shifted forward one letter in the alphabet, the result is IBM. Because the IBM logo is visible on HAL's display terminals, everybody assumed that the letter shift was intentional on Arthur Clarke's part. Clarke has since assured me that it was totally accidental, and that he was astounded when the shift was first called to his attention.