Original

book than a machine or when they otherwise resemble an information service, strict liability rules may not be imposed on them. As this article explains, courts have decided that books should not be treated as "products" for strict liability purposes and that publishers of books should not be held strictly liable in tort when their products contain defective information.

Remedies

When a seller has breached implied or express warranties in connection with the sale of goods, the buyer can sue the seller to recover money damages for certain kinds of injuries arising from the breach. If, for example, a consumer is physically injured by a defective lawnmower and has to pay $10,000 in medical expenses, that $10,000 may be recovered from the manufacturer or the firm from which the consumer bought the lawnmower. If the lawnmower must be repaired or replaced, the consumer can generally recover in contract for these damages as well.

Contract damages, however, tend to be more limited than tort damages. Monetary damages to compensate an injured person for pain and suffering, for example, are recoverable in tort actions (such as negligence and strict liability) but may not be in contract actions. Some economic losses are also not recoverable in contract cases. Unless, for example, the manufacturer (or other seller) of a lawnmower had reason to know at the time of the sale that a particular buyer of the lawnmower needed it to operate a lawn-mowing service, the buyer would not be able to recover lost profits on his lawn-mowing business during the time the business was out of operation after the defect in it evidenced itself.

In negligence actions, successful plaintiffs can generally recover damages for a broad range of injuries flowing from the negligent act, including pain and suffering and some economic losses. In strict liability actions, only damages arising from physical harms to persons or property are generally recoverable.

One other respect in which tort and contract actions tend to differ is in the kinds of persons who can bring claims for what kinds of damages. Contract law tends (except where physical injury to persons or property is involved) to limit the class of possible plaintiffs to those who bought the goods and are thus the beneficiaries of the warranty promises that are part of the contract. Tort law is more generous about who can bring a lawsuit (e.g., if the buyer of the product gives it to another person as a gift and that person is harmed, he or she can sue in tort whereas that person might not be able to sue in contract).

Multiple volumes of thick treatises have been written to explain all the nuances of contract and tort liability arising from defective products. This brief synopsis is necessarily incomplete but will, I hope, give those in the computing field some grounding in the basics of these legal categories.

REFERENCE

1. Arnold, R. The persistence of caveat emptor: Publisher immunity from liability for inaccurate factual information. *U. Pitt. Law Rev. 53* (Spring 1992), 777.

▷ ON THE IMPACT OF THE COMPUTER ON SOCIETY

How Does One Insult a Machine?

Joseph Weizenbaum

The structure of the typical essay on "The impact of computers on society" is as follows: First there is an "on the one hand" statement. It tells all the good things computers have already done for society and often even attempts to argue that the social order would already have collapsed were it not for the "computer revolution." This is usually followed by an "on the other hand" caution which tells of certain problems the introduction of computers brings in its wake. The threat posed to individual privacy by large data banks and the danger of large-scale unemployment induced by industrial automation are usually mentioned. Finally, the glorious present and prospective achievements of the computer are applauded, while the dangers alluded to in the second part are shown to be capable of being alleviated by sophisticated technological fixes. The closing paragraph consists of a plea for generous societal support for more, and more large-scale, computer research and development. This is usually coupled to the more or less subtle assertion that only computer science, hence only the computer scientist, can guard the world against the admittedly hazardous fallout of applied computer technology.

In fact, the computer has had very considerably less societal impact than the mass media would lead us to believe. Certainly, there are enterprises like space travel that could not have been undertaken without computers. Certainly the computer industry, and with it the computer education industry, has grown to enormous proportions. But much of the industry is self-serving. It is rather like an island economy in which the natives make a living by taking in each other's laundry. The part that is not self-serving is largely supported by government agencies and other gigantic enterprises that know the value of everything but the price of nothing, that is, that know the short-range utility of computer systems but have no idea of their ultimate social cost. In any case, airline reservation systems and computerized hospitals serve only a tiny, largely the most affluent, fraction of society. Such things cannot be said to have an impact on society generally.

SIDE EFFECTS OF TECHNOLOGY

The more important reason that I dismiss the argument which I have caricatured is that the direct societal effects of any pervasive new technology are as nothing compared to its much more subtle and ultimately much more important side effects. In that sense, the societal impact of the computer has not yet been felt.

To help firmly fix the idea of the importance of subtle indirect effects of technology, consider the impact on society of the invention of the microscope. When it was invented in the middle of the seventeenth century, the dominant commonsense theory of disease was fundamentally that disease was a punishment visited upon an individual by God. The sinner's body was thought to be inhabited by various so-called humors brought into disequilibrium in accordance with divine justice. The cure for disease was therefore to be found first in penance and second in the balancing of humors as, for example, by bleeding. Bleeding was, after all, both painful, hence punishment and penance, and potentially balancing in that it actually removed substance from the body. The microscope enabled man to see microorganisms and thus paved the way for the germ theory of disease. The enormously surprising discovery of extremely small living organisms also induced the idea of a continuous chain of life which, in turn, was a necessary intellectual precondition for the emergence of Darwinism. Both the germ theory of disease and the theory of evolution profoundly altered man's conception of his contract with God and consequently his self-image. Politically these ideas served to help diminish the power of the Church and, more generally, to legitimize the questioning of the basis of hitherto unchallenged authority. I do not say that the microscope alone was responsible for the enormous social changes that followed its invention. Only that it made possible the kind of paradigm shift, even on the commonsense level, without which these changes might have been impossible.

Is it reasonable to ask whether the computer will induce similar changes in man's image of himself and whether that influence will prove to be its most important effect on society? I think so, although I hasten to add that I don't believe the computer has yet told us much about man and his nature. To come to grips with the question, we must first ask in what way the computer is different from man's many other machines. Man has built two fundamentally different kinds of machines, nonautonomous and autonomous. An autonomous machine is one that operates for long periods of time, not on the basis of inputs from the real world, for example from sensors or from human drivers, but on the basis of internalized models of some aspect of the real world. Clocks are examples of autonomous machines in that they operate on the basis of an internalized model of the planetary system. The computer is, of course, the example par excellence. It is able to internalize models of essentially unlimited complexity and of a fidelity limited only by the genius of man.

It is the autonomy of the computer we value. When, for example, we speak of the power of computers as increasing with each new hardware and software development, we mean that, because of their increasing speed and storage capacity, and possibly thanks to new programming tricks, the new computers can internalize ever more complex and ever more faithful models of ever larger slices of reality. It seems strange then that, just when we exhibit virtually an idolatry of autonomy with respect to machines, serious thinkers in respected academies (I have in mind B. F. Skinner of Harvard University [1]) can rise to question autonomy as a fact for man. I do not think that the appearance of this paradox at this time is accidental. To understand it, we must realize that man's commitment to science has always had a masochistic component.

Time after time science has led us to insights that, at least when seen superficially, diminish man. Thus Galileo removed man from the center of the universe, Darwin removed him from his place separate from the animals, and Freud showed his rationality to be an illusion. Yet man pushes his inquiries further and deeper. I cannot help but think that there is an analogy between man's pursuit of scientific knowledge and an individual's commitment to psychoanalytic therapy. Both are undertaken in the full realization that what the inquirer may find may well damage his self-esteem. Both may reflect his determination to find meaning in his existence through struggle in truth, however painful that may be, rather than to live without meaning in a world of ill-disguised illusion. However, I am also aware that sometimes people enter psychoanalysis unwilling to put their illusions at risk, not searching for a deeper reality but in order to convert the insights they hope to gain to personal power. The analogy to man's pursuit of science does not break down with that observation.

Each time a scientific discovery shatters a hitherto fundamental cornerstone of the edifice on which man's self-esteem is built, there is an enormous reaction, just as is the case under similar circumstances in psychoanalytic therapy. Powerful defense mechanisms, beginning with denial and usually terminating in rationalization, are brought to bear. Indeed, the psychoanalyst suspects that, when a patient appears to accept a soul-shattering insight without resistance, his very casualness may well mask his refusal to allow that insight truly operational status in his self-image. But what is the psychoanalyst to think about the patient who positively embraces tentatively proffered, profoundly humiliating self-knowledge, when he embraces it and instantly converts it to a new foundation of his life? Surely such an event is symptomatic of a major crisis in the mental life of the patient.

I believe we are now at the beginning of just such a crisis in the mental life of our civilization. The microscope, I have argued, brought in its train a revision of man's image of himself. But no one in the mid-seventeenth century could have foreseen that. The possibility that the computer will, one way or another, demonstrate that, in the inimitable phrase of one of my esteemed colleagues, "the brain is merely a meat machine" is one that engages academicians, industrialists, and journalists in the here and now. How has the computer contributed to bringing about this very sad state of affairs? It must be said right away that the computer alone is not the chief causative agent. It is merely an extreme extrapolation of technology. When seen as an inducer of philosophical dogma, it is merely the reductio ad absurdum of a technological ideology. But how does it come to be regarded as a source of philosophic dogma?

THEORY VERSUS PERFORMANCE

We must be clear about the fact that a computer is nothing without a program. A program is fundamentally a transformation of one computer into another that has autonomy and that, in a very real sense, behaves. Programming languages describe dynamic processes. And, most importantly, the processes they describe can be actually carried out. Thus we can build models of any aspect of the real world that interests us and that we understand. And we can make our models work. But we must be careful to remember that a computer model is a description that works. Ordinarily, when we speak of A being a model of B, we mean that a theory about some aspects of the behavior of B is also a theory of the same aspects of the behavior of A. It follows that when, for example, we consider a computer model of paranoia, like that published by Colby *et al.* [2], we must not be persuaded that it tells us anything about paranoia on the grounds that it, in some sense, mirrors the behavior of a paranoiac. After all, a plain typewriter in some sense mirrors the behavior of an autistic child (one types a question and gets no response whatever), but it does not help us to understand autism. A model must be made to stand or fall on the basis of its theory. Thus, while programming languages may have put a new power in the hands of social scientists in that this new notation may have freed them from the vagueness of discursive descriptions, their obligation to build defensible theories is in no way diminished. Even errors can be pronounced with utmost formality and eloquence. But they are not thereby transmuted to truth.

The failure to make distinctions between descriptions, even those that "work," and theories accounts in large part for the fact that those who refuse to accept the view of man as machine have been put on the defensive. Recent advances in computer understanding of natural language offer an excellent case in point. Halle and Chomsky, to mention only the two with whom I am most familiar, have long labored on a theory of language which any model of language behavior must satisfy [3]. Their aim is like that of the physicist who writes a set of differential equations that anyone riding a bicycle must satisfy. No physicist claims that a person need know, let alone be able to solve, such differential equations in order to become a competent cyclist. Neither do Halle and Chomsky claim that humans know or knowingly obey the rules they believe to govern language behavior. Halle and Chomsky also strive, as do physical theorists, to identify the constants and parameters of their theories with components of reality. They hypothesize that their rules constitute a kind of projective description of certain aspects of the structure of the human mind. Their problem is thus not merely to discover economical rules to account for language behavior, but also to infer economic mechanisms which determine that precisely those rules are to be preferred over all others. Since they are in this way forced to attend to the human mind, not only that of speakers of English, they must necessarily be concerned with all human language behavior—not just that related to the understanding of English.

The enormous scope of their task is illustrated by their observation that in all human languages declarative sentences are often transformed into questions by a permutation of two of their words. (John is here → Is John here?) It is one thing to describe rules that transform declarative sentences into questions—a simple permutation rule is clearly insufficient—but another thing to describe a "machine" that necessitates those rules when others would, all else being equal, be simpler. Why, for example, is it not so that declarative sentences read backward transform those sentences into questions? The answer must be that other constraints on the "machine" combine against this local simplicity in favor of a more nearly global economy. Such examples illustrate the depth of the level of explanation that Halle and Chomsky are trying to achieve. No wonder that they stand in awe of their subject matter.

Workers in computer comprehension of natural language operate in what is usually called performance mode. It is as if they are building machines that can ride bicycles by following heuristics like "if you feel a displacement to the left, move your weight to the left." There can be, and often is, a strong interaction between the development of theory and the empirical task of engineering systems whose theory is not yet thoroughly understood. Witness the synergistic cooperation between aerodynamics and aircraft design in the first quarter of the present century. Still, what counts in performance mode is not the elaboration of theory but the performance of systems. And the systems being hammered together by the new crop of computer semanticists are beginning (just beginning) to perform.

Since computer scientists have recognized the importance of the interplay of syntax, semantics, and pragmatics, and with it the importance of computer-manipulable knowledge, they have made progress. Perhaps by the end of the present decade, computer systems will exist with which specialists, such as physicians and chemists and mathematicians, will converse in natural language. And surely some part of such achievements will have been based on other successes in, for example, computer simulation of cognitive processes. It is understandable that any success in this area, even if won empirically and without accompanying enrichments of theory, can easily lead to certain delusions being planted. Is it, after all, not terribly tempting to believe that a computer that understands natural language at all, however narrow the context, has captured something of the essence of man? Descartes himself might have believed it.

Indeed, by way of this very understandable seduction, the computer comes to be a source of philosophical dogma.

I am tempted to recite how performance programs are composed and how things that don't work quite correctly are made to work via all sorts of stratagems which do not even pretend to have any theoretical foundation. But the very asking of the question, "Has the computer captured the essence of man?" is a diversion and, in that sense, a trap. For the real question "Does man understand the essence of man?" cannot be answered by technology and hence certainly not by any technological instrument.

THE TECHNOLOGICAL METAPHOR

I asked earlier what the psychoanalyst is to think when a patient grasps a tentatively proffered deeply humiliating interpretation and attempts to convert it immediately to a new foundation of his life. I now think I phrased that question too weakly. What if the psychoanalyst merely coughed and the cough entrained the consequences of which I speak? That is our situation today. Computer science, particularly its artificial intelligence branch, has coughed. Perhaps the press has unduly amplified that cough—but it is only a cough nevertheless. I cannot help but think that the eagerness to believe that man's whole nature has suddenly been exposed by that cough, and that it has been shown to be a clockwork, is a symptom of something terribly wrong.

What is wrong, I think, is that we have permitted technological metaphors, what Mumford [4] calls the "Myth of the Machine," and technique itself to so thoroughly pervade our thought processes that we have finally abdicated to technology the very duty to formulate questions. Thus sensible men correctly perceive that large data banks and enormous networks of computers threaten man. But they leave it to technology to formulate the corresponding question. Where a simple man might ask: "Do we need these things?", technology asks "what electronic wizardry will make them safe?" Where a simple man will ask "is it good?", technology asks "will it work?" Thus science, even wisdom, becomes what technology and most of all computers can handle. Lest this be thought to be an exaggeration, I quote from the work of H. A. Simon, one of the most senior of American computer scientists [5]:

> As we succeed in broadening and deepening our knowledge—theoretical and empirical—about computers, we shall discover that in large part their behavior is governed by simple general laws, that what appeared as complexity in the computer program was, to a considerable extent, complexity of the environment to which the program was seeking to adapt its behavior.
>
> To the extent that this prospect can be realized, it opens up an exceedingly important role for computer simulation as a tool for achieving a deeper understanding of human behavior. For if it is the organization of components, and not their physical properties, that largely determines behavior, and if computers are organized somewhat in the image of man, then the computer becomes an obvious device for exploring the consequences of alternative organizational assumptions for human behavior.

and

> A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself.
>
> . . . I believe that this hypothesis holds even for the whole man.

We already know that those aspects of the behavior of computers which cannot be attributed to the complexity of their programs is governed by simple general laws—ultimately by the laws of Boolean algebra. And of course the physical properties of the computer's components are nearly irrelevant to its behavior. Mechanical relays are logically equivalent to tubes and to transistors and to artificial neurons. And of course the complexity of computer programs is due to the complexity of the environments, including the computing environments themselves, with which they were designed to deal. To what else could it possibly be due? So, what Simon sees as prospective is already realized. But does this collection of obvious and simple facts lead to the conclusion that man is as simple as are computers? When Simon leaps to that conclusion and then formulates the issue as he has done here, that is, when he suggests that the behavior of *the whole man* may be understood in terms of the behavior of computers as governed by simple general laws, then the very possibility of understanding man as an autonomous being, as an individual with deeply internalized values, that very possibility is excluded. How does one insult a machine?

The question "Is the brain merely a meat machine?", which Simon puts in a so much more sophisticated form, is typical of the kind of question formulated by, indeed formulatable only by, a technological mentality. Once it is accepted as legitimate, arguments as to what a computer can or cannot do "in principle" begin to rage and themselves become legitimate. But the legitimacy of the technological question—for example, is human behavior to be understood either in terms of the organization or of the physical properties of "components"—need not be admitted in the first instance. A human question can be asked instead. Indeed, we might begin by asking what has already become of "the whole man" when he can conceive of computers organized in his own image.

The success of technique and of some technological explanations has, as I've suggested, tricked us into permitting technology to formulate important questions for us—questions whose very forms severely diminish the number of degrees of freedom in our range of decision-making. Whoever dictates the questions in large part determines the answers. In that sense, technology, and especially computer technology, has become a self-fulfilling nightmare reminiscent of that of the lady who dreams of being raped and begs her attacker to be kind to her. He answers "it's your dream, lady." We must come to see that technology is our dream and that we must ultimately decide how it is to end.

I have suggested that the computer revolution need not and ought not to call man's dignity and autonomy into question, that it is a kind of pathology that moves men to wring from it unwarranted, enormously damaging interpretations. Is then the computer less threatening that we might have thought? Once we realize that our visions, possibly nightmarish visions, determine the effect of our own creations on us and on our society, their threat to us is surely diminished. But that is not to say that this realization alone will wipe out all danger. For example, apart from the erosive effect of a technological mentality on man's self-image, there are practical attacks on the freedom and dignity of man in which computer technology plays a critical role.

I mentioned earlier that computer science has come to recognize the importance of building knowledge into machines. We already have a machine—Dendral—[6] that commands more chemistry than do many Ph.D. chemists, and another—Mathlab—[7] that commands more applied mathematics than do many applied mathematicians. Both Dendral and Mathlab contain knowledge that can be evaluated in terms of the explicit theories from which it was derived. If the user believes that a result Mathlab delivers is wrong, then, apart from possible program errors, he must be in disagreement, not with the machine or its programmer, but with a specific mathematical theory. But what about the many programs on which management, most particularly the government and the military, rely, programs which can in no sense be said to rest on explicable theories but are instead enormous patchworks of programming techniques strung together to make them work?

INCOMPREHENSIBLE SYSTEMS

In our eagerness to exploit every advance in technique we quickly incorporate the lessons learned from machine manipulation of knowledge in theory-based systems into such patchworks. They then "work" better. I have in mind systems like target selection systems used in Vietnam and war games used in the Pentagon, and so on. These often gigantic systems are put together by teams of programmers, often working over a time span of many years. But by the time the systems come into use, most of the original programmers have left or turned their attention to other pursuits. It is precisely when gigantic systems begin to be used that their inner workings can no longer be understood by any single person or by a small team of individuals. Norbert Wiener, the father of cybernetics, foretold this phenomenon in a remarkably prescient article [8] published more than a decade ago. He said there:

> It may well be that in principle we cannot make any machine the elements of whose behavior we cannot comprehend sooner or later. This does not mean in any way that we shall be able to comprehend these elements in substantially less time than the time required for operation of the machine, or even within any given number of years or generations.
>
> An intelligent understanding of [machines'] mode of performance may be delayed until long after the task which they have been set has been completed. This means that though machines are theoretically subject to human criticism, such criticism may be ineffective until long after it is relevant.

This situation, which is now upon us, has two consequences: first that decisions are made on the basis of rules and criteria no one knows explicitly, and second that the system of rules and criteria becomes immune to change. This is so because, in the absence of detailed understanding of the inner workings of a system, any substantial modification is very likely to render the system altogether inoperable. The threshold of complexity beyond which this phenomenon occurs has already been crossed by many existing systems, including some compiling and computer operating systems. For example, no one likes the operating systems for certain large computers, but they cannot be substantially changed nor can they be done away with. Too many people have become dependent on them.

An awkward operating system is inconvenient. That is not too bad. But the growing reliance on supersystems that were perhaps designed to help people make analyses and decisions, but which have since surpassed the understanding of their users while at the same time becoming indispensable to them, is another matter. In modern war it is common for the soldier, say the bomber pilot, to operate at an enormous psychological distance from his victims. He is not responsible for burned children because he never sees their village, his bombs, and certainly not the flaming children themselves. Modern technological rationalizations of war, diplomacy, politics, and commerce such as computer games have an even more insidious effect on the making of policy. Not only have policy makers abdicated their decision-making responsibility to a technology they don't understand, all the while maintaining the illusion that they, the policy makers, are formulating policy questions and answering them, but responsibility has altogether evaporated. No human is any longer responsible for "what the machine says." Thus there can be neither right nor wrong, no question of justice, no theory with which one can agree or disagree, and finally no basis on which one can challenge "what the machine says." My father used to invoke the ultimate authority by saying to me, "it is written." But then I could read what was written, imagine a human author, infer his values, and finally agree or disagree. The systems in the Pentagon, and their counterparts elsewhere in our culture, have in a very real sense no authors. They therefore do not admit of exercises of imagination that may ultimately lead to human judgment. No wonder that men who live day in and out with such machines and become dependent on them begin to believe that men are merely machines. They are reflecting what they themselves have become.

The potentially tragic impact on society that may ensue from the use of systems such as I have just discussed is greater than might at first be imagined. Again it is side effects, not direct effects, that matter most. First, of course, there is the psychological impact on individuals living in a society in which anonymous, hence irresponsible, forces formulate the large questions of the day and circumscribe the range of possible answers. It cannot be surprising that large numbers of perceptive individuals living in such a society experience a kind of impotence and fall victim to the mindless rage that often accompanies such experiences. But even worse, since computer-based knowledge systems become essentially unmodifiable except in that they can grow, and since they induce dependence and cannot, after a certain threshold is crossed, be abandoned, there is an enormous risk that they will be passed from one generation to another, always growing. Man too passes knowledge from one generation to another. But because man is mortal, his transmission of knowledge over the generations is at once a process of filtering and accrual. Man doesn't merely pass knowledge, he rather regenerates it continuously. Much as we may mourn the crumbling of ancient civilizations, we know nevertheless that the glory of man resides as much in the evolution of his cultures as in that of his brain. The unwise use of ever larger and ever more complex computer systems may well bring this process to a halt. It could well replace the ebb and flow of culture with a world without values, a world in which what counts for a fact has long ago been determined and forever fixed.

POSITIVE EFFECTS

I've spoken of some potentially dangerous effects of present computing trends. Is there nothing positive to be said? Yes, but it must be said with caution. Again, side effects are more important than direct effects. In particular, the idea of computation and of programming languages is beginning to become an important metaphor which, in the long run, may well prove to be responsible for paradigm shifts in many fields. Most of the common-sense paradigms in terms of which much of mankind interprets the phenomena of the everyday world, both physical and social, are still deeply rooted in fundamentally mechanistic metaphors. Marx's dynamics as well as those of Freud are, for example, basically equilibrium systems. Any hydrodynamicist could come to understand them without leaving the jargon of his field. Languages capable of describing ongoing processes, particularly in terms of modular subprocesses, have already had an enormous effect on the way computer people think of every aspect of their worlds, not merely those directly related to their work. The information-processing view of the world so engendered qualifies as a genuine metaphor. This is attested to by the fact that it (i) constitutes an intellectual framework that permits new questions to be asked about a wide-ranging set of phenomena, and (ii) that it itself provides criteria for the adequacy of proffered answers. A new metaphor is important not in that it may be better than existing ones, but rather in that it may enlarge man's vision by giving him yet another perspective on his world. Indeed, the very effectiveness of a new metaphor may seduce lazy minds to adopt it as a basis for universal explanations and as a source of panaceas. Computer simulation of social processes has already been advanced by single-minded generalists as leading to general solutions of all of mankind's problems.

The metaphors given us by religion, the poets, and by thinkers like Darwin, Newton, Freud, and Einstein have rather quickly penetrated to the language of ordinary people. These metaphors have thus been instrumental in shaping our entire civilization's imaginative reconstruction of our world. The computing metaphor is as yet available to only an extremely small set of people. Its acquisition and internalization, hopefully as only one of many ways to see the world, seems to require experience in program composition, a kind of computing literacy. Perhaps such literacy will become very widespread in the advanced societal sectors of the advanced countries. But, should it become a dominant mode of thinking and be restricted to certain social classes, it will prove not merely repressive in the ordinary sense, but an enormously divisive societal force. For then classes which do and do not have access to the metaphor will, in an important sense, lose their ability to communicate with one another. We know already how difficult it is for the poor and the oppressed to communicate with the rest of the society in which they are embedded. We know how difficult it is for the world of science to communicate with that of the arts and of the humanities. In both instances the communication difficulties, which have grave consequences, are very largely due to the fact that the respective communities have unsharable experiences out of which unsharable metaphors have grown.

RESPONSIBILITY

Given these dismal possibilities, what is the responsibility of the computer scientist? First I should say that most of the harm computers can potentially entrain is much more a function of properties people attribute to computers than of what a computer can or cannot actually be made to do. The nonprofessional has little choice but to make his attributions of properties to computers on the basis of the propaganda emanating from the computer community and amplified by the press. The computer professional therefore has an enormously important responsibility to be modest in his claims. This advice would not even have to be voiced if computer science had a tradition of scholarship and of self-criticism such as that which characterizes the established sciences. The mature scientist stands in awe before the depth of his subject matter. His very humility is the wellspring of his strength. I regard the instilling of just this kind of humility, chiefly by the example set by teachers, to be one of the most important missions of every university department of computer science.

The computer scientist must be aware constantly that his instruments are capable of having gigantic direct and indirect amplifying effects. An error in a program, for example, could have grievous direct results, including most certainly the loss of much human life. On 11 September 1971, to cite just one example, a computer programming error caused the simultaneous destruction of 117 high-altitude weather balloons whose instruments were being monitored by an earth satellite [9]. A similar error in a military command and control system could launch a fleet of nuclear tipped missiles. Only censorship prevents us from knowing how many such events involving non-nuclear weapons have already occurred. Clearly then, the computer scientist has a heavy responsibility to make the fallibility and limitations of the systems he is capable of designing brilliantly clear. The very power of his systems should serve to inhibit the advice he is ready to give and to constrain the range of work he is willing to undertake.

Of course, the computer scientist, like everyone else, is responsible for his actions and their consequences. Sometimes that responsibility is hard to accept because the corresponding authority to decide what is and what is not to be done appears to rest with distant and anonymous forces. That technology itself determines what is to be done by a process of extrapolation and that individuals are powerless to intervene in that determination is precisely the kind of self-fulfilling dream from which we must awaken.

Consider gigantic computer systems. They are, of course, natural extrapolations of the large systems we already have. Computer networks are another point on the same curve extrapolated once more. One may ask whether such systems can be used by anybody except by governments and very large corporations and whether such organizations will not use them mainly for antihuman purposes. Or consider speech recognition systems. Will they not be used primarily to spy on private communications? To answer such questions by saying that big computer systems, computer networks, and speech recognition systems are inevitable is to surrender one's humanity. For such an answer must be based either on one's profound conviction that society has already lost control over its technology or on the thoroughly immoral position that "if I don't do it, someone else will."

I don't say that systems such as I have mentioned are necessarily evil—only that they may be and, what is most important, that their inevitability cannot be accepted by individuals claiming autonomy, freedom, and dignity. The individual computer scientist can and must decide. The determination of what the impact of computers on society is to be is, at least in part, in his hands.

Finally, the fundamental question the computer scientist must ask himself is the one that every scientist, indeed every human, must ask. It is not "what shall I do?" but rather "what shall I be?" I cannot answer that for anyone save myself. But I will say again that if technology is a nightmare that appears to have its own inevitable logic, it is our nightmare. It is possible, given courage and insight, for man to deny technology the prerogative to formulate man's questions. It is possible to ask human questions and to find humane answers.

NOTES

1. B. F. Skinner, *Beyond Freedom and Dignity* (Knopf, New York, 1971).
2. K. M. Colby, S. Weber, F. D. Hilf, *Artif. Intell.* 1, 1 (1971).
3. N. Chomsky, *Aspects of the Theory of Syntax* (M.I.T. Press, Cambridge, Mass., 1965); ——— and M. Halle, *The Sound Pattern of English* (Harper & Row, New York, 1968).

4. L. Mumford, *The Pentagon of Power* (Harcourt, Brace, Jovanovich, New York, 1970).
5. H. A. Simon, *The Sciences of the Artificial* (M.I.T. Press, Cambridge, Mass., 1969), pp. 22–25.
6. B. Buchanan, G. Sutherland, and E. A. Feigenbaum, in *Machine Intelligence*, B. Meltzer, Ed. (American Elsevier, New York, 1969).
7. W. A. Martin and R. J. Fateman, "The Macsyma system," in *Proceedings of the 2nd Symposium on Symbolic and Algebraic Manipulation* (Association for Computer Machinery, New York, 1971); J. Moses, *Commun. Assoc. Computer Mach.* 14 (No. 8), 548 (1971).
8. N. Wiener, *Science* 131, 1355 (1960).
9. R. Gillette, *ibid.* 174, 477 (1971).

▷ PROFESSIONAL ETHICS

Deborah G. Johnson

SCENARIO ONE: CONFLICTING LOYALTIES

Carl Babbage is an experienced systems designer. He has been working for the Acme Software Company for over three years. Acme develops and sells computer hardware and software. It does this both by designing and marketing general purpose systems and by contracting with companies and government agencies to design systems for their exclusive use.

During the first two years that Carl worked for Acme, he worked on software that Acme was developing for general marketing. A year ago, however, he was reassigned to work on a project under contract with the U.S. Defense Department. The project involves designing a system that will monitor radar signals and launch nuclear missiles in response to these signals.

Carl initially had some reluctance about working on a military project, but he put this out of his mind because the project seemed challenging and he knew that if he did not work on it, someone else would. Now, however, the project is approaching completion and Carl has some grave reservations about the adequacy of the system. He is doubtful about the system's capacity for making fine distinctions (for example, distinguishing between a small aircraft and a missile) and the security of the mechanism that can launch missiles (for example, it may be possible for unauthorized individuals to get access to the controls under certain circumstances). Carl expressed his concern to the project director but she dismissed these concerns quickly, mentioning that Acme is already behind schedule on the project and has already exceeded the budget that they had agreed to with the Defense Department.

Carl feels that he has a moral responsibility to do something, but he doesn't know what to do. Should he ask for reassignment to another project? Should he go