JOSEPH WEIZENBAUM

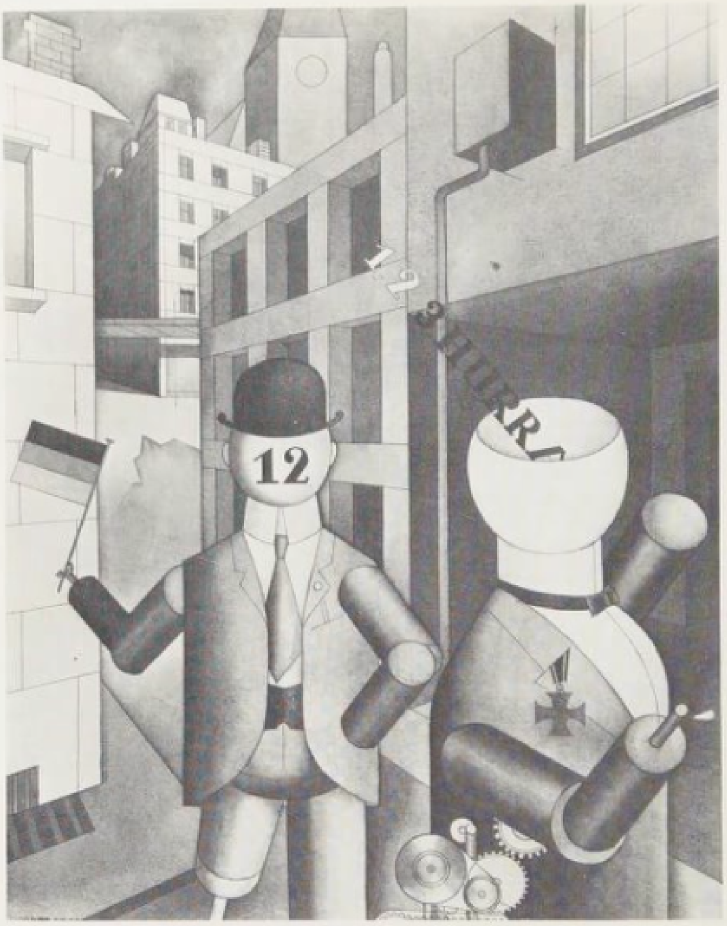
COMPUTER POWER AND HUMAN REASON

FROM JUDGMENT TO CALCULATION

Flip right

Tom
Cooper

COMPUTER POWER
AND
HUMAN REASON



Republican Automatons (1920) by George Grosz.
Watercolor, 23 5/8 x 18 5/8". Collection, The Museum
of Modern Art, New York, Advisory Committee Fund.

COMPUTER POWER
AND
HUMAN REASON

FROM JUDGMENT TO CALCULATION

Joseph Weizenbaum
THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

W. H. FREEMAN AND COMPANY
San Francisco

Library of Congress Cataloging in Publication Data

Weizenbaum, Joseph.
Computer power and human reason.

Includes bibliographical references.
1. Computers. 2. Computers and civilization.
3. Programming (Electronic computers) I. Title.
QA76.W44 001.6'4 75-19305
ISBN 0-7167-0464-1
ISBN 0-7167-0463-3 pbk.

Printed in the United States of America

9 8 7

to Ruth

CONTENTS

PREFACE ix

INTRODUCTION 1

1. ON TOOLS 17
2. WHERE THE POWER OF THE COMPUTER COMES FROM 39
3. HOW COMPUTERS WORK 73
4. SCIENCE AND THE COMPULSIVE PROGRAMMER 111
5. THEORIES AND MODELS 132
6. COMPUTER MODELS IN PSYCHOLOGY 154
7. THE COMPUTER AND NATURAL LANGUAGE 182
8. ARTIFICIAL INTELLIGENCE 202
9. INCOMPREHENSIBLE PROGRAMS 228
10. AGAINST THE IMPERIALISM OF INSTRUMENTAL REASON 258

NOTES 281

INDEX 289

PREFACE

This book is only nominally about computers. In an important sense, the computer is used here merely as a vehicle for moving certain ideas that are much more important than computers. The reader who looks at a few of this book's pages and turns away in fright because he spots an equation or a bit of computer jargon here and there should reconsider. He may think that he does not know anything about computers, indeed, that computers are too complicated for ordinary people to understand. But a major point of this book is precisely that we, all of us, have made the world too much into a computer, and that this remaking of the world in the image of the computer started long before there were any electronic computers. Now that we have computers, it becomes somewhat easier to see this imaginative transformation we have worked on the world. Now we can use the computer itself—that is the idea of the computer—as a metaphor to help us understand what we have done and are doing.

We are all used to hearing that the computer is a powerful new instrument. But few people have any idea where the power of a computer comes from. Chapters I to III are devoted to explaining just that. With a modest investment in time and intellectual energy, anyone who can read this Preface should be able to work his way through those chapters. Chapters II and III will be the most difficult, but the reader who cannot master them should not therefore abandon the rest of the book. Really, the only point Chapters II and III make is that computers are in some sense "universal" machines, that they can (in a certain sense which is there explained) do "anything." The reader who is willing to take that assertion on faith may well wish to skip from Chapter I (which he should read) to Chapter IV. Perhaps after he has finished the whole book, he will be tempted to try Chapters II and III again.

The rest of the book contains the major arguments, which are in essence, first, that there is a difference between man and machine, and, second, that there are certain tasks which computers *ought* not be made to do, independent of whether computers *can* be made to do them.

The writing of this book has been an adventure to me. First and most important, I have been cheered beyond my power to say by the generosity and the intellectual and emotional support given me by people who owe me absolutely nothing. But now I am very greatly in their debt. I am thinking primarily of Lewis Mumford, that grand old man, of Noam Chomsky, and of Steven Marcus, the literary critic. Each of them read large sections of the manuscript in preparation (Lewis Mumford read all of it) and contributed the wisest and most useful kinds of criticism. But more than that, each encouraged me to go on when I despaired. For there was often cause for despair. I am acutely aware, for example, that there is nothing I say in this book that has not been said better, certainly more eloquently, by others. But, as my friends continued to point out to me, it seemed important to say these things again and again. And, as Lewis Mumford often remarked, it sometimes matters that a member of the scientific establishment say some things that humanists have been shouting for ages.

More formally, I am indebted to my University, the Massachusetts Institute of Technology, for granting me a leave of two

years so that I might pursue first the thinking which preceded the writing, then the writing itself. I spent the first of those two happy years at the Center for Advanced Study in the Behavioral Sciences at Stanford, California.* It was there that I met Steven Marcus, as well as others of the Center's Fellows, who struggled mightily to educate a primitive engineer. The names John Platt, Paul Armer, Herbert Weiner, Fredrick Redlich, Alexander Mitscherlich, and Israel Scheffler immediately come to mind. I thank them for their efforts. No failure of mine should, however, be counted against any of them. I spent the second year as a Vinton Hayes Research Fellow at Harvard University. There I had the good fortune to be able to renew an old collegial association, namely, one with Professor Thomas Cheatham, an outstanding computer scientist. He took the trouble to read almost the whole manuscript as it sprang from my pen. Professor Hilary Putnam of Harvard's Philosophy Department gave me many hours of his valuable time. Without his help, encouragement, and guidance I would have fallen into many more traps than I actually did. It was also a stroke of good luck that Daniel C. Dennett, an outstanding young philosopher from Tufts University, happened to be spending the year at Harvard just when I was there. His patience with my philosophical naivety was unlimited. I can never adequately discharge my debts to all these good people.

These few words of thanks acknowledge the fact that this book—like, I suspect, most others—has many co-authors whose names will not appear on its cover. But in this instance that confession would be grievously incomplete if it did not include an acknowledgment of the critical contributions that the book's manuscript editor, Aidan Kelly, made to it. I cannot, in a few words, summarize what he did. Perhaps readers will understand if I say simply that Aidan Kelly is a poet.

Finally, everyone who has ever written a book will know what an enormous burden such a task imposes on the author's family. My wife, Ruth, suffered my retreats to my study with the utmost

* My fellowship was supported in part by National Science Foundation Grant No. SSH71-01834 A01 from the Research Applied to National Needs (RANN) Program of NSF to the Center. Of course, the opinions, findings, conclusions, and recommendations contained herein are entirely mine and do not necessarily reflect the views of any sponsor.

good will and patience. She helped me over the inevitable bouts with the feelings of guilt that overcome an author when he is writing—for then he is not with his family even when he is with them—and when he is not writing—for then he is not doing what he has set himself to do. My children counted the pages as they mounted on my desk. And they grieved when, as often happened, the stack of pages in the wastebasket grew more quickly than that on my desk. Most of all, they cheerfully endured the endless progress reports that punctuated our dinner-table conversation. This book is Ruth's and our children's as much as it is mine.

Fall 1975 *Joseph Weizenbaum*
Cambridge, Massachusetts

COMPUTER POWER
AND
HUMAN REASON

INTRODUCTION

In 1935, Michael Polanyi, then holder of the Chair of Physical Chemistry at the Victoria University of Manchester, England, was suddenly shocked into a confrontation with philosophical questions that have ever since dominated his life. The shock was administered by Nicolai Bukharin, one of the leading theoreticians of the Russian Communist party, who told Polanyi that "under socialism the conception of science pursued for its own sake would disappear, for the interests of scientists would spontaneously turn to the problems of the current Five Year Plan." Polanyi sensed then that "the scientific outlook appeared to have produced a mechanical conception of man and history in which there was no place for science itself." And further that "this conception denied altogether any intrinsic power to thought and thus denied any grounds for claiming freedom of thought."[1]

I don't know how much time Polanyi thought he would devote to developing an argument for a contrary concept of man and history. His very shock testifies to the fact that he was in profound disagreement with Bukharin, therefore that he already conceived of man differently, even if he could not then give explicit form to his concept. It may be that he determined to write a counterargument to Bukharin's position, drawing only on his own experience as a scientist, and to have done with it in short order. As it turned out, however, the confrontation with philosophy triggered by Bukharin's revelation was to demand Polanyi's entire attention from then to the present day.

I recite this bit of history for two reasons. The first is to illustrate that ideas which seem at first glance to be obvious and simple, and which ought therefore to be universally credible once they have been articulated, are sometimes buoys marking out stormy channels in deep intellectual seas. That science is creative, that the creative act in science is equivalent to the creative act in art, that creation springs only from autonomous individuals, is such a simple and, one might think, obvious idea. Yet Polanyi has, as have many others, spent nearly a lifetime exploring the ground in which it is anchored and the turbulent sea of implications which surrounds it.

The second reason I recite this history is that I feel myself to be reliving part of it. My own shock was administered not by any important political figure espousing his philosophy of science, but by some people who insisted on misinterpreting a piece of work I had done. I write this without bitterness and certainly not in a defensive mood. Indeed, the interpretations I have in mind tended, if anything, to overrate what little I had accomplished and certainly its importance. No, I recall that piece of work now only because it seems to me to provide the most parsimonious way of identifying the issues I mean to discuss.

The work was done in the period 1964–1966, and was reported in the computer-science literature in January 1966 and August 1967. To summarize it briefly, I composed a computer program with which one could "converse" in English. The human conversationalist partner would type his portion of the conversation on a typewriter connected to a computer, and the computer, under control of my program, would analyze the message that had so been transmitted to it, compose a response to it in English, and cause the response to be typed on the computer's typewriter.

I chose the name ELIZA for the language analysis program because, like the Eliza of Pygmalion fame, it could be taught to "speak" increasingly well. Because conversations must be about something, that is, because they must take place within some context, the program was constructed in a two-tier arrangement, the first tier consisting of the language analyzer and the second of a script. The script is a set of rules rather like those that might be given to an actor who is to use them to improvise around a certain theme. Thus ELIZA could be given a script to enable it to maintain a conversation about cooking eggs or about managing a bank checking account, and so on. Each specific script thus enabled ELIZA to play a specific conversational role.

For my first experiment, I gave ELIZA a script designed to permit it to play (I should really say parody) the role of a Rogerian psychotherapist engaged in an initial interview with a patient. The Rogerian psychotherapist is relatively easy to imitate because much of his technique consists of drawing his patient out by reflecting the patient's statements back to him. The following conversation between a young lady and Eliza playing doctor illustrates both the Rogerian technique of encouraging a patient to keep talking and the operation of the computer program ELIZA. The first to "speak" is the young lady. The computer's responses are printed entirely in capitals.[2]

Men are all alike.

IN WHAT WAY

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

It's true. I am unhappy.

DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

I need some help, that much seems certain.

WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

Perhaps I could learn to get along with my mother.

TELL ME MORE ABOUT YOUR FAMILY

My mother takes care of me.

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU

My father.

YOUR FATHER

You are like my father in some ways.

WHAT RESEMBLANCE DO YOU SEE

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

WHAT MAKE YOU THINK I AM NOT VERY AGGRESSIVE

You don't argue with me.

WHY DO YOU THINK I DON'T ARGUE WITH YOU

You are afraid of me.

DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

My father is afraid of everybody.

WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER

Bullies.

DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

DOCTOR, as ELIZA playing psychiatrist came to be known, soon became famous around the Massachusetts Institute of Technology, where it first came into existence, mainly because it was an easy program to demonstrate. Most other programs could not vividly demonstrate the information-processing power of a computer to visitors who did not already have some specialized knowledge, say,

of some branch of mathematics. DOCTOR, on the other hand, could be appreciated on some level by anyone. Its power as a demonstration vehicle was further enhanced by the fact that the visitor could actually participate in its operation. Soon copies of DOCTOR, constructed on the basis of my published description of it, began appearing at other institutions in the United States. The program became nationally known and even, in certain circles, a national plaything.

The shocks I experienced as DOCTOR became widely known and "played" were due principally to three distinct events.

1. A number of practicing psychiatrists seriously believed the DOCTOR computer program could grow into a nearly completely automatic form of psychotherapy. Colby *et al.* write, for example,

> "Further work must be done before the program will be ready for clinical use. If the method proves beneficial, then it would provide a therapeutic tool which can be made widely available to mental hospitals and psychiatric centers suffering a shortage of therapists. Because of the time-sharing capabilities of modern and future computers, several hundred patients an hour could be handled by a computer system designed for this purpose. The human therapist, involved in the design and operation of this system, would not be replaced, but would become a much more efficient man since his efforts would no longer be limited to the one-to-one patient-therapist ratio as now exists."[3]*

I had thought it essential, as a prerequisite to the very possibility that one person might help another learn to cope with his emotional problems, that the helper himself participate in the other's experience of those problems and, in large part by way of his own em-

* Nor is Dr. Colby alone in his enthusiasm for computer administered psychotherapy. Dr. Carl Sagan, the astrophysicist, recently commented on ELIZA in *Natural History*, vol. LXXXIV, no. 1 (Jan. 1975), p. 10: "No such computer program is adequate for psychiatric use today, but the same can be remarked about some human psychotherapists. In a period when more and more people in our society seem to be in need of psychiatric counseling, and when time sharing of computers is widespread, I can imagine the development of a network of computer psychotherapeutic terminals, something like arrays of large telephone booths, in which, for a few dollars a session, we would be able to talk with an attentive, tested, and largely non-directive psychotherapist."

pathic recognition of them, himself come to understand them. There are undoubtedly many techniques to facilitate the therapist's imaginative projection into the patient's inner life. But that it was possible for even one practicing psychiatrist to advocate that this crucial component of the therapeutic process be entirely supplanted by pure technique—*that* I had not imagined! What must a psychiatrist who makes such a suggestion think he is doing while treating a patient, that he can view the simplest mechanical parody of a single interviewing technique as having captured anything of the essence of a human encounter? Perhaps Colby *et al.* give us the required clue when they write;

> "A human therapist can be viewed as an information processor and decision maker with a set of decision rules which are closely linked to short-range and long-range goals, . . . He is guided in these decisions by rough empiric rules telling him what is appropriate to say and not to say in certain contexts. To incorporate these processes, to the degree possessed by a human therapist, in the program would be a considerable undertaking, but we are attempting to move in this direction."[4]

What can the psychiatrist's image of his patient be when he sees himself, as therapist, not as an engaged human being acting as a healer, but as an information processor following rules, etc.?

Such questions were my awakening to what Polanyi had earlier called a "scientific outlook that appeared to have produced a mechanical conception of man."

2. I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchanges with it, she asked me to leave the room. Another time, I suggested I might rig the system so that I could examine all conversations anyone had had with it, say, overnight. I was promptly bombarded with accusations that what I proposed amounted to spying on people's most

intimate thoughts; clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms. I knew of course that people form all sorts of emotional bonds to machines, for example, to musical instruments, motorcycles, and cars. And I knew from long experience that the strong emotional ties many programmers have to their computers are often formed after only short exposures to their machines. What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people. This insight led me to attach new importance to questions of the relationship between the individual and the computer, and hence to resolve to think about them.

3. Another widespread, and to me surprising, reaction to the ELIZA program was the spread of a belief that it demonstrated a general solution to the problem of computer understanding of natural language. In my paper, I had tried to say that no general solution to that problem was possible, i.e., that language is understood only in contextual frameworks, that even these can be shared by people to only a limited extent, and that consequently even people are not embodiments of any such general solution. But these conclusions were often ignored. In any case, ELIZA was such a small and simple step. Its contribution was, if any at all, only to vividly underline what many others had long ago discovered, namely, the importance of context to language understanding. The subsequent, much more elegant, and surely more important work of Winograd[5] in computer comprehension of English is currently being misinterpreted just as ELIZA was. This reaction to ELIZA showed me more vividly than anything I had seen hitherto the enormously exaggerated attributions an even well-educated audience is capable of making, even strives to make, to a technology it does not understand. Surely, I thought, decisions made by the general public about emergent technologies depend much more on what that public attributes to such technologies than on what they actually are or can and cannot do. If, as appeared to be the case, the public's attributions are wildly misconceived, then public decisions are bound to be misguided and

often wrong. Difficult questions arise out of these observations; what, for example, are the scientist's responsibilities with respect to making his work public? And to whom (or what) is the scientist responsible?

As perceptions of these kinds began to reverberate in me, I thought, as perhaps Polanyi did after his encounter with Bukharin, that the questions and misgivings that had so forcefully presented themselves to me could be disposed of quickly, perhaps in a short, serious article. I did in fact write a paper touching on many points mentioned here.* But gradually I began to see that certain quite fundamental questions had infected me more chronically than I had first perceived. I shall probably never be rid of them.

There are as many ways to state these basic questions as there are starting points for coping with them. At bottom they are about nothing less than man's place in the universe. But I am professionally trained only in computer science, which is to say (in all seriousness) that I am extremely poorly educated; I can mount neither the competence, nor the courage, not even the chutzpah, to write on the grand scale actually demanded. I therefore grapple with questions that couple more directly to the concerns I have expressed, and hope that their larger implications will emerge spontaneously.

I shall thus have to concern myself with the following kinds of questions:

1. What is it about the computer that has brought the view of man as a machine to a new level of plausibility? Clearly there have been other machines that imitated man in various ways, e.g., steam shovels. But not until the invention of the digital computer have there been machines that could perform intellectual functions of even modest scope; i.e., machines that could in any sense be said to be intelligent. Now "artificial intelligence" (AI) is a subdiscipline of computer science. This new field will have to be discussed. Ultimately a line dividing human and machine intelligence must be drawn. If there is no such line, then advocates of computerized psychotherapy may be merely heralds of an age in which man has finally been recognized as nothing but a clock-work. Then the consequences of such a reality would need urgently to be divined and contemplated.

2. The fact that individuals bind themselves with strong emotional ties to machines ought not in itself to be surprising. The instruments man uses become, after all, extensions of his body. Most importantly, man must, in order to operate his instruments skillfully, internalize aspects of them in the form of kinesthetic and perceptual habits. In that sense at least, his instruments become literally part of him and modify him, and thus alter the basis of his affective relationship to himself. One would expect man to cathect more intensely to instruments that couple directly to his own intellectual, cognitive, and emotive functions than to machines that merely extend the power of his muscles. Western man's entire milieu is now pervaded by complex technological extensions of his every functional capacity. Being the enormously adaptive animal he is, man has been able to accept as authentically natural (that is, as given by nature) such technological bases for his relationship to himself, for his identity. Perhaps this helps to explain why he does not question the appropriateness of investing his most private feelings in a computer. But then, such an explanation would also suggest that the computing machine represents merely an extreme extrapolation of a much more general technological usurpation of man's capacity to act as an autonomous agent in giving meaning to his world. It is therefore important to inquire into the wider senses in which man has come to yield his own autonomy to a world viewed as machine.

3. It is perhaps paradoxical that just, when in the deepest sense man has ceased to believe in—let alone to trust—his own autonomy, he has begun to rely on autonomous machines, that is, on machines that operate for long periods of time entirely on the basis of their own internal realities. If his reliance on such machines is to be based on something other than unmitigated despair or blind faith, he must explain to himself what these machines do and even how they do what they do. This requires him to build some conception of their internal "realities." Yet most men don't understand computers to even the slightest degree. So, unless they are capable of very great skepticism (the kind we bring to bear while watching a stage magi-

cian), they can explain the computer's intellectual feats only by bringing to bear the single analogy available to them, that is, their model of their own capacity to think. No wonder, then, that they overshoot the mark; it is truly impossible to imagine a human who could imitate ELIZA, for example, but for whom ELIZA's language abilities were his limit. Again, the computing machine is merely an extreme example of a much more general phenomenon. Even the breadth of connotation intended in the ordinary usage of the word "machine," large as it is, is insufficient to suggest its true generality. For today when we speak of, for example, bureaucracy, or the university, or almost any social or political construct, the image we generate is all too often that of an autonomous machine-like process.

These, then, are the thoughts and questions which have refused to leave me since the deeper significances of the reactions to ELIZA I have described began to become clear to me. Yet I doubt that they could have impressed themselves on me as they did were it not that I was (and am still) deeply involved in a concentrate of technological society as a teacher in the temple of technology that is the Massachusetts Institute of Technology, an institution that proudly boasts of being "polarized around science and technology." There I live and work with colleagues, many of whom trust only modern science to deliver reliable knowledge of the world. I confer with them on research proposals to be made to government agencies, especially to the Department of "Defense." Sometimes I become more than a little frightened as I contemplate what we lead ourselves to propose, as well as the nature of the arguments we construct to support our proposals. Then, too, I am constantly confronted by students, some of whom have already rejected all ways but the scientific to come to know the world, and who seek only a deeper, more dogmatic indoctrination in that faith (although that word is no longer in their vocabulary). Other students suspect that not even the entire collection of machines and instruments at M.I.T. can significantly help give meaning to their lives. They sense the presence of a dilemma in an education polarized around science and technology, an education that implicitly claims to open a privileged

access-path to fact, but that cannot tell them how to decide what is to count as fact. Even while they recognize the genuine importance of learning their craft, they rebel at working on projects that appear to address themselves neither to answering interesting questions of fact nor to solving problems in theory.

Such confrontations with my own day-to-day social reality have gradually convinced me that my experience with ELIZA was symptomatic of deeper problems. The time would come, I was sure, when I would no longer be able to participate in research proposal conferences, or honestly respond to my students' need for therapy (yes, that is the correct word), without first attempting to make sense of the picture my own experience with computers had so sharply drawn for me.

Of course, the introduction of computers into our already highly technological society has, as I will try to show, merely reinforced and amplified those antecedent pressures that have driven man to an ever more highly rationalistic view of his society and an ever more mechanistic image of himself. It is therefore important that I construct my discussion of the impact of the computer on man and his society so that it can be seen as a particular kind of encoding of a much larger impact, namely, that on man's role in the face of technologies and techniques he may not be able to understand and control. Conversations around that theme have been going on for a long time. And they have intensified in the last few years.

Certain individuals of quite differing minds, temperaments, interests, and training have—however much they differ among themselves and even disagree on many vital questions—over the years expressed grave concern about the conditions created by the unfettered march of science and technology; among them are Mumford, Arendt, Ellul, Roszak, Comfort, and Boulding. The computer began to be mentioned in such discussions only recently. Now there are signs that a full-scale debate about the computer is developing. The contestants on one side are those who, briefly stated, believe computers can, should, and will do everything, and on the other side those who, like myself, believe there are limits to what computers ought to be put to do.

It may appear at first glance that this is an in-house debate of little consequence except to a small group of computer technicians. But at bottom, no matter how it may be disguised by technological jargon, the question is whether or not every aspect of human thought is reducible to a logical formalism, or, to put it into the modern idiom, whether or not human thought is entirely computable. That question has, in one form or another, engaged thinkers in all ages. Man has always striven for principles that could organize and give sense and meaning to his existence. But before modern science fathered the technologies that reified and concretized its otherwise abstract systems, the systems of thought that defined man's place in the universe were fundamentally juridicial. They served to define man's obligations to his fellow men and to nature. The Judaic tradition, for example, rests on the idea of a contractual relationship between God and man. This relationship must and does leave room for autonomy for both God and man, for a contract is an agreement willingly entered into by parties who are free not to agree. Man's autonomy and his corresponding responsibility is a central issue of all religious systems. The spiritual cosmologies engendered by modern science, on the other hand, are infected with the germ of logical necessity. They, except in the hands of the wisest scientists and philosophers, no longer content themselves with explanations of appearances, but claim to say how things actually are and must necessarily be. In short, they convert truth to provability.

As one consequence of this drive of modern science, the question, "What aspects of life are formalizable?" has been transformed from the moral question, "How and in what form may man's obligations and responsibilities be known?" to the question, "Of what technological genus is man a species?" Even some philosophers whose every instinct rebels against the idea that man is entirely comprehensible as a machine have succumbed to this spirit of the times. Hubert Dreyfus, for example, trains the heavy guns of phenomenology on the computer model of man.[7] But he limits his argument to the technical question of what computers can and cannot do. I would argue that if computers could imitate man in every respect—which in fact they cannot—even then it would be appropriate, nay, urgent, to examine the computer in the light of man's perennial need to find his place in the world. The outcomes of practical matters that are of vital importance to everyone hinge on how and in what terms the discussion is carried out.

One position I mean to argue appears deceptively obvious: it is simply that there are important differences between men and machines as thinkers. I would argue that, however intelligent machines may be made to be, there are some acts of thought that *ought* to be attempted only by humans. One socially significant question I thus intend to raise is over the proper place of computers in the social order. But, as we shall see, the issue transcends computers in that it must ultimately deal with logicality itself—quite apart from whether logicality is encoded in computer programs or not.

The lay reader may be forgiven for being more than slightly incredulous that anyone should maintain that human thought is entirely computable. But his very incredulity may itself be a sign of how marvelously subtly and seductively modern science has come to influence man's imaginative construction of reality.

Surely, much of what we today regard as good and useful, as well as much of what we would call knowledge and wisdom, we owe to science. But science may also be seen as an addictive drug. Not only has our unbounded feeding on science caused us to become dependent on it, but, as happens with many other drugs taken in increasing dosages, science has been gradually converted into a slow-acting poison. Beginning perhaps with Francis Bacon's misreading of the genuine promise of science, man has been seduced into wishing and working for the establishment of an age of rationality, but with his vision of rationality tragically twisted so as to equate it with logicality. Thus have we very nearly come to the point where almost every genuine human dilemma is seen as a mere paradox, as a merely apparent contradiction that could be untangled by judicious applications of cold logic derived from a higher standpoint. Even murderous wars have come to be perceived as mere problems to be solved by hordes of professional problemsolvers. As Hannah Arendt said about recent makers and executors of policy in the Pentagon:

> "They were not just intelligent, but prided themselves on being 'rational' . . . They were eager to find formulas, preferably expressed in a pseudo-mathematical language, that would unify the

> most disparate phenomena with which reality presented them; that is, they were eager to discover *laws* by which to explain and predict political and historical facts as though they were as necessary, and thus as reliable, as the physicists once believed natural phenomena to be . . . [They] did not *judge;* they calculated. . . . an utterly irrational confidence in the calculability of reality [became] the leitmotif of the decision making."[8]

And so too have nearly all political confrontations, such as those between races and those between the governed and their governors, come to be perceived as mere failures of communication. Such rips in the social fabric can then be systematically repaired by the expert application of the latest information-handling techniques—at least so it is believed. And so the rationality-is-logicality equation, which the very success of science has drugged us into adopting as virtually an axiom, has led us to deny the very existence of human conflict, hence the very possibility of the collision of genuinely incommensurable human interests and of disparate human values, hence the existence of human values themselves.

It may be that human values are illusory, as indeed B. F. Skinner argues. If they are, then it is presumably up to science to demonstrate that fact, as indeed Skinner (as scientist) attempts to do. But then science must itself be an illusory system. For the only certain knowledge science can give us is knowledge of the behavior of formal systems, that is, systems that are games invented by man himself and in which to assert truth is nothing more or less than to assert that, as in a chess game, a particular board position was arrived at by a sequence of legal moves. When science purports to make statements about man's experiences, it bases them on identifications between the primitive (that is, undefined) objects of one of its formalisms, the pieces of one of its games, and some set of human observations. No such sets of correspondences can ever be proved to be correct. At best, they can be falsified, in the sense that formal manipulations of a system's symbols may lead to symbolic configurations which, when read in the light of the set of correspondences in question, yield interpretations contrary to empirically observed phenomena. Hence all empirical science is an elaborate structure built on piles that are anchored, not on bedrock as is commonly

supposed, but on the shifting sand of fallible human judgment, conjecture, and intuition. It is not even true, again contrary to common belief, that a single purported counter-instance that, if accepted as genuine would certainly falsify a specific scientific theory, generally leads to the immediate abandonment of that theory. Probably all scientific theories currently accepted by scientists themselves (excepting only those purely formal theories claiming no relation to the empirical world) are today confronted with contradicting evidence of more than negligible weight that, again if fully credited, would logically invalidate them. Such evidence is often explained (that is, explained away) by ascribing it to error of some kind, say, observational error, or by characterizing it as inessential, or by the assumption (that is, the faith) that some yet-to-be-discovered way of dealing with it will some day permit it to be acknowledged but nevertheless incorporated into the scientific theories it was originally thought to contradict. In this way scientists continue to rely on already impaired theories and to infer "scientific fact" from them.*

The man in the street surely believes such scientific facts to be as well-established, as well-proven, as his own existence. His certitude is an illusion. Nor is the scientist himself immune to the same illusion. In his praxis, he must, after all, suspend disbelief in order to do or think anything at all. He is rather like a theatergoer, who, in order to participate in and understand what is happening on the stage, must for a time pretend to himself that he is witnessing real events. The scientist must believe his working hypothesis, together with its vast underlying structure of theories and assumptions, even if only for the sake of the argument. Often the "argument" extends over his entire lifetime. Gradually he becomes what he at first merely pretended to be: a true believer. I choose the word "argument" thoughtfully, for scientific demonstrations, even mathematical proofs, are fundamentally acts of persuasion.

* Thus, Charles Everett writes on the now-discarded phlogiston theory of combustion (in the *Encyclopaedia Britannica*, 11th ed., 1911, vol. VI, p. 34): "The objections of the antiphlogistonists, such as the fact that the calices weigh more than the original metals instead of less as the theory suggests, were answered by postulating that phlogiston was a principle of levity, or even completely ignored as an accident, the change in qualities being regarded as the only matter of importance." Everett lists H. Cavendish and J. Priestley, both great scientists of their time, as adherents to the phlogiston theory.

Scientific statements can never be certain; they can be only more or less credible. And credibility is a term in individual psychology, i.e., a term that has meaning only with respect to an individual observer. To say that some proposition is credible is, after all, to say that it is believed by an agent who is free not to believe it, that is, by an observer who, after exercising judgment and (possibly) intuition, chooses to accept the proposition as worthy of his believing it. How then can science, which itself surely and ultimately rests on vast arrays of human value judgments, demonstrate that human value judgments are illusory? It cannot do so without forfeiting its own status as the single legitimate path to understanding man and his world.

But no merely logical argument, no matter how cogent or eloquent, can undo this reality: that science has become the sole legitimate form of understanding in the common wisdom. When I say that science has been gradually converted into a slow-acting poison, I mean that the attribution of certainty to scientific knowledge by the common wisdom, an attribution now made so nearly universally that it has become a commonsense dogma, has virtually delegitimatized all other ways of understanding. People viewed the arts, especially literature, as sources of intellectual nourishment and understanding, but today the arts are perceived largely as entertainments. The ancient Greek and Oriental theaters, the Shakespearian stage, the stages peopled by the Ibsens and Chekhovs nearer to our day—these were schools. The curricula they taught were vehicles for understanding the societies they represented. Today, although an occasional Arthur Miller or Edward Albee survives and is permitted to teach on the New York or London stage, the people hunger only for what is represented to them to be scientifically validated knowledge. They seek to satiate themselves at such scientific cafeterias as *Psychology Today*, or on popularized versions of the works of Masters and Johnson, or on scientology as revealed by L. Ron Hubbard. Belief in the rationality-logicality equation has corroded the prophetic power of language itself. We can count, but we are rapidly forgetting how to say what is worth counting and why.

1

ON TOOLS

The stories of man and of his machines are inseparably woven together. Machines have enabled man to transform his physical environment. With their aid he has plowed the land and built cities and dug great canals. These transformations of man's habitat have necessarily induced mutations in his societal arrangements. But even more crucially, the machines of man have strongly determined his very understanding of his world and hence of himself. Man is conscious of himself, of the existence of others like himself, and of a world that is, at least to some extent, malleable. Most importantly, man can foresee. In the act of designing implements to harrow the pliant soil, he rehearses their action in his imagination. Moreover, since he is conscious of himself as a social creature and as one who will inevitably die, he is necessarily a teacher. His tools, whatever their primary practical function, are necessarily also pedagogical instruments. They are then part of the stuff out of which man fashions his imaginative reconstruction of the world. It is within the intellectual and social world he himself creates that the individual prehearses and rehearses countless dramatic enactments of how the world might have been and what it might become. That world is the repository of his subjectivity. Therefore it is the stimulator of his consciousness and finally the constructor of the material world itself. It is this self-constructed world that the individual encounters as an apparently external force. But he contains it within himself; what confronts him is his own model of a universe, and, since he is part of it, his model of himself.

Man can create little without first imagining that he can create it. We can imagine the rehearsal of how he would use it that must have gone on in a stone-age man while he laboriously constructed his axe. Did not each of us recapitulate this ancestral experience when as small children we constructed primitive toys of whatever material lay within our reach? But tools and machines do not merely signify man's imaginativeness and its creative reach, and they are certainly not important merely as instruments for the transformation of a malleable earth: they are pregnant symbols in themselves. They symbolize the activities they enable, i.e., their own use. An oar is a tool for rowing, and it represents the skill of rowing in its whole complexity. No one who has not rowed can see an oar as truly an oar. The way someone who has never played one sees the violin is simply not the same, by very far, as the way a violinist sees it. A tool is also a model for its own reproduction and a script for the reenactment of the skill it symbolizes. That is the sense in which it is a pedagogic instrument, a vehicle for instructing men in other times and places in culturally acquired modes of thought and action. The tool as symbol in all these respects thus transcends its role as a practical means toward certain ends: it is a constituent of man's symbolic recreation of his world. It must therefore inevitably enter into the imaginative calculus that constantly constructs his world. In that sense, then, the tool is much more than a mere device: it is an agent for change. It is even more than a fragment of a blueprint of a world determined for man and bequeathed to him by his forebearers—although it is that too.

It is readily understandable that hand-held tools and especially hand-held weapons have direct effects on the imaginations of individuals who use them. When hunters acquired spears, for example, they must have seen themselves in an entirely new relationship to their world. Large animals which had earlier raided their foodstores and even attacked their children and which they feared, now became man's prey. Man's source of food grew, for now men could kill animals at a distance, including many species that had eluded them before. The effectively greater abundance of food must also have enlarged the domain over which they could range, thus increasing the likelihood that they would meet other people. Their experience of the world changed and so too must have their idea of their place in it.

The six-shooter of the nineteenth-century American West was known as the "great equalizer," a name that eloquently testifies to what that piece of hardware did to the self-image of gun-toters who, when denuded of their weapons, felt themselves disadvantaged with respect to their fellow citizens. But devices and machines, perhaps known to (and certainly owned and operated by) only a relatively few members of society, have often influenced the self-image of its individual members and the world-image of the society as a whole quite as profoundly as have widely used hand tools. Ships of all kinds, for example, were instrumental in informing man of the vastness of his domain. They permitted different cultures to meet and to cross-fertilize one another. The seafarer's ships and all his other artifacts, his myths and legends, effectively transmitted his lore from generation to generation. And they informed the unconscious of those who stayed on the land as much as that of those who actually sailed. The printing press transformed the world even for those millions who, say, in Martin Luther's time, remained illiterate and perhaps never actually saw a book. And of the great masses of people all over the world whose lives were directly and dramatically changed by the industrial revolution, how many ever actually operated a steam engine? Nor is modern society immune to huge shocks administered as side effects of the introduction of new machines. The cotton-picking machine was deployed in the cotton fields of the American South beginning about 1955. It quickly destroyed the market for the only thing vast masses of black Southern agricultural workers had to sell: their labor. Thus began the mass migration of the American Black to the cities, particularly to such northern manufacturing centers as Detroit, Chicago, and New York, but also to the large Southern cities, such as Birmingham and Atlanta. Surely this enormous change in the demography of the United States, this internal migration of millions of its citizens, was and remains one of the principal determinants of the course of the American civil-rights movement. And that movement has nontrivially influenced the consciousness of every American at least, if not of almost every living adult anywhere on this earth.

What is the compelling urgency of the machine that it can so intrude itself into the very stuff out of which man builds his world?

Many machines are functional additions to the human body, virtually prostheses. Some, like the lever and the steam shovel extend the raw muscular power of their individual operators; some, like the microscope, the telescope, and various measuring instruments, are extensions of man's sensory apparatus. Others extend the physical reach of man. The spear and the radio, for example, permit man to cast his influence over a range exceeding that of his arms and voice, respectively. Man's vehicles make it possible for him to travel faster and farther than his legs alone would carry him, and they allow him to transport great loads over vast distances. It is easy to see how and why such prosthetic machines directly enhance man's sense of power over the material world. And they have an important psychological effect as well: they tell man that he can remake himself. Indeed, they are part of the set of symbols man uses to recreate his past, i.e., to construct his history, and to create his future. They signify that man, the engineer, can transcend limitations imposed on him by the puniness of his body and of his senses. Once man could kill another animal only by crushing or tearing it with his hands; then he acquired the axe, the spear, the arrow, the ball fired from a gun, the explosive shell. Now charges mounted on missiles can destroy mankind itself. That is one measure of how far man has extended and remade himself since he began to make tools.

To construe the influence of prosthetic tools on man's transformation entirely in terms of the power they permitted man to

aggregate to himself may invite a view of man's relationship to nature whose principal—indeed, almost sole—component is a raw power struggle. Man, in this view, finally conquered nature simply by mustering sufficient power to overcome natural space and time, to engineer life and death, and finally to destroy the earth altogether. But this idea is mistaken, even if we accept that man's eternal dream has been, not merely the discovery of nature, but its conquest, and that that dream has now been largely realized. For if victory over nature has been achieved in this age, then the nature over which modern man reigns is a very different nature from that in which man lived before the scientific revolution. Indeed, the trick that man turned and that enabled the rise of modern science was nothing less than the transformation of nature and of man's perception of reality.

The paramount change that took place in the mental life of man, beginning during roughly the fourteenth century, was in man's perception of time and consequently of space. Man had long ago noticed (and, we may suppose, thought about) regularities in the world about him. Alexander Marshack has shown that even Upper Paleolithic man (circa 30,000 B.C.) had a notation for lunar cycles that was, in Marshack's words, "already evolved, complex and sophisticated, a tradition that would seem to have been thousands of years old by this point."[1] But from Classical antiquity until relatively recently, the regularity of the universe was searched for and perceived in great thematic harmonies. The idea that nature behaves systematically in the sense we understand it—i.e., that every part and aspect of nature may be isolated as a subsystem governed by laws describable as functions of time—this idea could not have been even understood by people who perceived time, not as a collection of abstract units (i.e., hours, minutes, and seconds), but as a sequence of constantly recurring events.

Times of day were known by events, such as the sun standing above a specific pile of rocks, or, as Homer tells us, by tasks begun or ended, such as the yoking of the oxen (morning) and the unyoking of the oxen (evening). Durations were indicated by reference to common tasks, e.g., the time needed to travel a well-known distance or to boil fixed quantities of water. Seasonal times were known by recurring seasonal events, e.g., the departure of birds.

Until Darwin's theory of evolution began to sink into the stream of commonly held ideas, i.e., to become "common sense," people knew that the world about them—the world of reproducing plants and animals, of rivers that flowed and dried up and flowed again, of seas that pulsed in great tidal rhythms, and of the ever-repeating spectacles in the heavens—had always existed, and that its fundamental law was eternal periodicity. Cosmological time, as well as the time perceived in daily life, was therefore a sort of complex beating, a repeating and echoing of events. Perhaps we can vaguely understand it by contemplating, say, the great fugues of Bach. But a special form of contemplation is required of us: we must not think in the modern manner, i.e., of Bach as a "problem solver," or of each of his *opera* in his *Art of the Fugue* as being his increasingly refined "solution" to a problem he had originally set himself. Instead we must think that Bach had the whole plan in his mind all the time, that he thought of the *Art of the Fugue* as a unified work with no beginning and no end, itself eternal like the cosmos, and like it enormously intricate in its connections, circles within circles within circles. We might then find it possible to think of life as having been not merely punctuated but entirely suffused by this kind of music, both on the grand cosmological-theological scale and on the small day-to-day level. Such time is a revolution of cycles and epicycles within cycles, not the receptacle of a uniformly flowing progression of abstract moments we now "know" it to be. And nature itself consisted, to be sure, of individual phenomena, but individual phenomena that were constantly repeating metamorphoses of themselves, and hence were permanent, eternal. "What is eternal is circular, and what is circular is eternal," Aristotle said, and even Galileo still believed the universe to be eternal and to be governed by recurrence and periodicity.

Darwin's understanding of time was radically different. He saw nature itself as a process in time and the individual phenomena of nature as irreversible metamorphoses. But he was far from being the originator of the idea of progress that is now so much with us. Indeed, he would not have been able to think his thoughts at all, if something very nearly like our current ideas of time had not already been an integral part of the common sense of his era.

How man's perception of time changed from that of the ancients to ours illuminates the role played by another kind of machine (one that is not prosthetic) in man's transformation from a creature of and living in nature to nature's master.

The clock is not a prosthetic machine; its product is not an extension of man's muscles or senses, but hours, minutes, and seconds, and today even micro-, nano-, and pico-seconds. Lewis Mumford calls the clock, not the steam engine, "the key machine of the modern industrial age."[2] In the brilliant opening chapter of his *Technics and Civilization*, he describes, among other things, how during the Middle Ages the ordered life of the monasteries affected life in the communities adjacent to them.

> "The monastery was the seat of a regular life. . . . the habit of order itself and the earnest regulation of time-sequences had become almost second nature in the monastery. . . . the monasteries—at one time there were 40,000 under the Benedictine rule—helped to give human enterprise the regular collective beat and rhythm of the machine; for the clock is not merely a means of keeping track of the hours, but of synchronizing the actions of men. . . . by the thirteenth century there are definite records of mechanical clocks, and by 1370 a well-designed 'modern' clock had been built by Heinrich von Wyck at Paris. Meanwhile, bell towers had come into existence, and the new clocks, if they did not have, till the fourteenth century, a dial and a hand that translated the movement of time into a movement through space, at all events struck the hours. The clouds that could paralyze the sundial . . . were no longer obstacles to time-keeping: summer or winter, day or night, one was aware of the measured clank of the clock. The instrument presently spread outside the monastery; and the regular striking of the bells brought a new regularity into the life of the workman and the merchant. The bells of the clock tower almost defined urban existence. Time-keeping passed into time-serving and time-accounting and time-rationing. As this took place, Eternity ceased gradually to serve as the measure and focus of human actions."[3]

Mumford goes on to make the crucial observation that the clock "disassociated time from human events and helped create the belief in an independent world of mathematically measurable sequences: the special world of science."[4] The importance of that effect of the clock on man's perception of the world can hardly be exagger-