JOSEPH WEIZENBAUM

COMPUTER POWER AND HUMAN REASON

FROM JUDGMENT TO CALCULATION
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."
Introduction

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
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,
tion 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."*

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."