DOCUMENT RESUME

ED 040 416 CG 005 401

AUTHOR Smith, David F.
TITLE Practical Consideration in the Application of Data Processing Techniques to the Needs of the Psychologist.
INSTITUTION American Personnel and Guidance Association, Washington, D.C.; Kentucky Univ., Lexington. Dept. of Educational Psychology.
PUB DATE 24 Mar 70
NOTE 6p.; Paper presented at American Personnel and Guidance Association Convention, New Orleans, Louisiana, March 22-26, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$0.40
DESCRIPTORS *Computer Oriented Programs, *Computer Programs, Computers, Computer Science, Computer Storage Devices, *Man Machine Systems, Psychology

ABSTRACT

Past problems encountered in the use of computers are discussed. They include: (1) computing and data storage costs; (2) machine speed and capacity limitations; and (3) the limited number of languages and Input/Output devices available to facilitate man/machine communication. The improvements in all of these areas are explained. The problems which remain, it was contended, lie in the area of communication. For example, the desires and demands of a programmer and a psychologist and the expectations of the computer are not in concert. Data rates for all three are discussed in order to emphasize the magnitude of the discrepancy between what men and machines can do. The paper concludes with the modest hope that soon a machine can be made to communicate only with the scope and clarity of an intelligent man. (TL)

Practical Considerations in the Application of Data Processing Techniques to the Needs of The Psychologist

David F. Smith, M.B.A.
Department of Educational Psychology
University of Kentucky
Lexington, Kentucky 40506

Read at the Annual Meeting
of
The American Personnel and Guidance Association
in
New Orleans, Louisiana
March 24, 1970

I. Problems-- Real or Imaginary?

Ten years ago, economic consideration of computing costs, data storage costs, and machine speed and capacity limitations would have discouraged application of the computer to anything more than the most basic and highest priority usages for counselors and psychologists.

Furthermore, only a limited number of languages and I/O devices were available to facilitate man/machine communications. These languages (and devices) were oriented towards mathematical operations and fixed-attribute data.

Happily, most of the machine-based problems-- with one big exception-- are vanishing. Central Processor costs have dropped to the point that time on a large-scale computer-- for instance an IBM 360/65 with 750,000 bytes (characters) of high speed core and 1 million bytes of moderate speed bulk core-- costs perhaps $2.00 to $5.00 per minute with access to a wide ranges of fast input/output devices. An on-line file of 12 million characters of high speed direct access costs only about $800.00 for the storage media and a small monthly charge for residence or mounting devices.

Devices and prototype machines now appearing in the data processing literature promise to boost speeds, enlarge storages, and decrease costs by approximately another order of magnitude in the next 5 years, or so. Examples of those devices are found in Large Scale Integrated circuits and memories,

holographic memories, non-mechanical high-speed printers, and scanning devices.

Typically, if cards are used for input, we may expect that a dollar spent on input will cost us a cent to process. This is a gross generalization which we'll dwell upon in a minute.

Languages have made great progress during this same period. At the point in time mentioned, you might have had a choice of perhaps three languages on a fairly large machine of the day; all oriented toward mathematics and all very close to bring mnemonic characterizations of the machine's "natural" language.

Today's languages are characterized by string handling facilities and logical data manipulation in addition to mathematic capability. Good examples are SNOBOL and LISP. Furthermore, modern languages almost invariably include capability to extend the language by use of the language for any special properties desired by the applications programmer. This is done by defining macro language, functions, or subroutines within the context of the language itself. There are perhaps 200 such languages extant-- signifying that no one language is universal to date, but the basic features desired abound. A large data-processing installation today usually offers 10 to 20 language support encompassing almost all of the features available in any computer language.

Several languages, notably Barnett's SNAP, offer "computer programming in English"; and, to a degree, succeed.

II. The Real Problem

Where, then, are the problems? in a word, the problems lie in the area of communication. Show me a Programmer and a Psychologist at one another's

throats and I'll show you three entities attempting to communicate-- the men and the machine.

Turing's test for intelligence has been passed by the machine. A recent (and humorous) example was quoted by Ellis (1959).

In all probability, the psychologist or counselor only desires that some benign agency under his direction and tutelage perform observations upon a large number of individuals over a "sufficient" period of time (whatever that is!) and do some combination of the following: (a) analyze the individual's differences from and similarity to several populations in view of diverse circumstances, (b) predict all reasonable outcomes (with probabilities) of the individual's present and future actions in light of the above, with and without various forms of intervention. (c) Prescribe various courses of action for the psychologist or counselor, or his underlings with due reference to Parkinson's and Peter's theories of management. (With appropriate warnings.) (d) Produce annually (and/or on short notice) professionally sound research papers to be read before the A.P.G.A. and/or A.P.A.-- as appropriate, and (e) be instantly recognized as so indubitably germane and useful and low-cost as to assure the financial security of the psychologist or counselor.

The programmer, in like fashion, only demands that the data be structured, complete, machine readable, self-indexed and, if possible, serial. Of course, he expects the same attributes to apply to the rules described by the psychologist or counselor. In addition, the rules must come in units easily codeable between coffee breaks.

The machine expects only the asynchronous arrival of heterogenous data.

The purpose of this levity is to help impress you with the volume of data that is implied when the professional comes to the programmer and says, "I want to develop a system to counsel school children, based on data to be collected from a series of interviews. (professional points proudly to a stack of video tape containing interviews)."

To put the problem in perspective from the machines point of view, we must consider the fact that the machines are performing operations at picosecond rates. Since a picosecond is one one-millionth of a millionth of a second, the machines have theoretical bandwidths measured in gigacycles per second. Contrast this with the bandwidth of a color television picture-- i.e. six megacycles per second or less. Clearly we cannot over-run the bigger machines processing rates. (We should note that the large machines almost always operate in some mode that alleviates this bottle-neck, be it multiprogramming, time slicing, spooling, or whatever).

Now, lets look at data rates for the Programmer and the Psychologist. An exceptionally productive programmer might produce debugged programs at a rate of ten to twenty instructions per hour. Reduced to machine terms, this might be 4000 8-bit characters per hour or 32,000 bits. That's a bit rate of approximately 10 cycles per second. Hardly picosecond rates.

The Psychologist might be able to outline his problems to the programmer at rates of a thousand times greater. If that rate prevails, then he lags the machine by only a factor of 10^8.

Problems of this magnitude defy easy solution. First steps may lie in elimination of the programmer as we now him. This first gain of 10^3 in rate might come about through efforts now being made in computational linguistics.

An excellant summary of the state of the art is to be found in "Natural Language Question-Answering Systems: 1969," Robert F. Simmons, CACM (January 1970) 15-30.

We have penetrated the problem of natural language-- machine communication just far enough to become aware of the scope of the problem, not far enough to begin to solve it.

The remaining lag of 10^8 bits per second is anyone's guess. Recent work in prostheses for the blind suggests that some direct form of man-machine thought coupling is possible, as science fiction has long predicted. I will be amazed if within my lifetime a machine can be made to communicate in natural language only with the scope and clarity of an intelligent man. Should direct thought communication become possible, then we shall have reached a state where not only the programmer vanishes, but perhpas also the counselor and the psychologist.