[*Times Literary Supplement*: Freeing the mind, April1962]

6—THE INTELLECT'S NEW EYE

By Margaret Masterman

IN THE OTHER ARTICLES in this series the digital computer has been thought of as a purely menial tool—in fact, as a kind of intellectual spade. This, it has been shown, can indeed assist a human scholar who is working on language, by performing for him a series of irksome repetitive tasks: tasks such as counting the number of words in a text, making a concordance of their occurrence, and mechanically reversing a large bi-lingual dictionary. It is, however, characteristic of such tasks that, when he programmes them on the computer, the scholar already knows—all too well—what kind of thing the output will be; he also knows exactly what increase of knowledge (or of doubt) will be the outcome of performing them. It is true that their performance depends on the computer in the sense that the scholar, unaided, just cannot get through them. They take too long, they are backbreaking, they are eye-wearing, they strain too far human capacity for maintaining accuracy; in fact, they are both physically and intellectually crushing. But they provoke no new theoretic vision, intuitive or mathematical. In this article I wish to suggest another quite different use for the digital computer in this non-numerical, data-processing field, namely, its potential use not as a tool but as a telescope. Of course, in one sense, a telescope is a tool, just as a spade is; but in another important sense it is not. For the use of a telescope, especially after its great development in the seventeenth century, so enlarged the whole range of what its possessors could see and do that, in the end, it was a factor in changing their whole picture of the world. I suppose that a spade which was sufficiently powerful to cut rapidly through the earth's outer crust—as the deep-boring drills of the future probably will—would be a tool in this new sense—that is, if its use reliably transformed our whole picture of the interior of the earth. But an ordinary spade would not, since it has not the right order of power. The potential capacity of the digital computer to process non-numerical data in novel ways—that capacity the surface of which has hardly been scratched as yet—is so great as to make of it the telescope of the mind.

There is a sentence in Leibniz in which he says that Wilkins's "Universal Character", that universal information-classifying system of which the thinkers of the seventeenth century all dreamed, would exalt the intellect (I misquote slightly) as much as Mr. Newton's telescope had exalted the vision. This prophecy was genuine, but its application misdirected. In my view, its fulfilment has come only now, three centuries later, in the development of the digital computer. For this, taken as a system, can indeed be envisaged, if you look at it from this point of view, as the universal mould into which any new form of knowledge must be poured.

* * *

As an example of this new power, and following up Leibniz's thought, consider in general the activity of classification. Long before the seventeenth century, and long after, thinkers ranging from Aristotle to Nelson Goodman, interested in exploring the general nature of classifying, had not yet succeeded in obtaining any new fundamental knowledge of it. Now, however, using digital computers, a new and elegant mathematical theory of classification is being developed all over the world. To get the feel of this, consider the following problem: You have discovered 150 characteristics, characterizing 100 so-called "species" (say of tapioca plants, or diseases, or sorts of words, or what-have-you). The problem is to find out how these newly discovered characteristics re-group up the 100 antecedently named species. Seen in terms of the groupings of these 150 characteristics, for instance, have we really got 100 species of tapioca plant? If not, how many have we? Human intuition, unaided, fails here; for we have too much classifying material. By making a 100 x150 array, however, into the squares of which we put a 1 wherever a species has one of the characteristics, we can provide basic long-range data for a computer. And by adopting some agreed-to-be-satisfactory criterion of similarity between any new species, we can provide it also with its immediate fodder. (Tanimoto's criterion of such similarity, for instance, is the number of properties in common between the two species divided by the number shown by at least one of them—so that, if the two species had three properties in common, the first having five altogether and the second six, the coefficient would be 318.)

This preparation achieved, a whole new classificatory countryside opens out; for, using the computer, the measures of similarity thus obtained can be compared, ordered and clustered, and the resulting clusters or clumps recompared with the original data, which may then suddenly appear in a new light. The range of application of this new kind of analysis is clearly enormous; already it is being, or has been, applied to classificatory problems in information retrieval, linguistics, medicine and anthropology. But that's only the start of the potential range of application; and that's only the start of the development of the theory. For, once you have begun to think in this new kind of way, it becomes clear that other non-obvious criteria of similarity can be tried out with sometimes the most unexpected effects.

* * *

This can be seen by repeating an experiment recently done by hand by two Cambridge undergraduates, who, classificatorily speaking, could not let well alone. Take two copies of each of the six volumes of the *Language Through Pictures* paperbacks, each of which has four stylized-stick-pictures per page. Cut each page in four, with a printer's guillotine. Stick the resulting pile of pictures on cards, number these and shuffle; you will now have a shuffled pack of over a thousand cards, each card bearing a stylized ideographic picture of a human situation. Sort these cards into two piles, in any apparently trivial way which occurs to you; e.g., sort all the pictures in which the sun appears from the pictures in which it does not. Record the sort, by noting the numbers of one pile of the cards; and then search for other defining characteristics of the subpacks you have obtained. In an astonishing number of cases you will find them, even when the sorting principle appeared ridiculous in the first place.

For instance, when the experiment was first performed, it was discovered that the presence of a table in a picture was of deep significance; in this world, pictures with tables represented homes. Moreover, the presence of a hat hanging on a hat-stand was a sign of continuity with earlier pictures of a hatted-man: it meant (in that world) Hat-Man, i.e., "That Man"; and so on, and so on. From this it was only a step to start identifying, interpreting and interrelating sequences of sorts; *for the pack of cards was now suddenly seen as a language.* The experiment finally foundered on two attempts (unsuccessful in both cases) to make the system simulate the ontological proof of the Existence of God, and also to make it describe the structure of itself. The point is, in the course of performing this experiment, the experimenters' whole vision of the experimental material had changed.

Try another vision-changing, data-processing game which does not require a machine. This game has two players. Take two copies of the English Penguin edition of *Roget's Thesaurus,* and give one copy to each player. The first player opens the *Thesaurus* at random, and with his eyes shut picks a *topic* (a topic is a headed paragraph) which is to be the goal of that play. The other player similarly opens the *Thesaurus* at random, and, with his eyes shut, picks a topic which is to be the start of that play. The object of the game is to move as rapidly as possible from some word in the starting topic to some word in the goal topic, *using only the cross-references in Roget's Thesaurus.* Each player moves in turn, moves are not disclosed, and one cross-reference hop from topic to topic counts as one move. It is said (I have not verified this) that a skilled player can get in not more than three moves from any topic to any other in *Roget's Thesaurus.*

* * *

"Yes, interesting", you will say, "but illuminating, why?" Consider, however, the revolutionary conception of language which is brought to light by the fact that this game can be played with *Roget's Thesaurus.* For Roget has classified the word-uses of the English language (not the words merely, but their uses) not alphabetically, nor yet according to their parts of speech, but according to the topics to which they can be applied; that is, according to their basic contexts. Given these contexts, his cross-reference system then defines language as a web (as it is suggested that language should be defined, in the last article in this series) in which the points of the web are the topics, and the cross-references the lines between them.

But who is it who for centuries has suggested that language should be defined like this? Not the practical men, the scientists; but the arts men, the philosophers and poets. St. Augustine, with his language of Platonic names (if that's what he meant); Owen Barfield, with his primitive set of archetypal poetic meanings; their intuition can now be analysed with a computer. And already the fact that this can be done has produced an illumination. For this game was devised by A. Newell and H. A. Simon (of the Carnegie Institute of Technology), whose main work is to make a computer throw light on the way human beings do mathematical proofs, by treating mathematical proofs as, in certain respects, analogous to the cross-referenced association-chains of Roget, used in the game. Moreover, computers, programmed in this way, can prove theorems; by using such methods, and, incredible as it may sound all these centuries after, a new proof of a theorem of Euclid has been obtained.

Moreover, using *Roget's Thesaurus* to illustrate the mechanics of the translation process, a new translation of the "omnis" in "Gallia est omnis divisa in partes tres" has been obtained. To see this, and get the "feel" of the whole process, a new game, though this time more a game of the patience type and also using *Roget's Thesaurus*, has to be played. This game, however, is frustrating in that it hardly ever works, it being indeed the aim of those who streamline the size of thesauruses for publication to ensure that this very game can be played as few times as possible. In order to play it, at any rate for a first attempt (that is, until you have become both skilled and brazen at adding strings of words to the original *Roget),* the use of the exact set of figures given below is advised. The object of play is the comparative elucidation of the sense of the word "point" in the following sentences, using exact methods only:

The last speaker made a number of facetious points.

The point of the steeple was blown off in the gale.

The point of the project is to investigate the nature of chromosomes.

The point of his argument was to persuade her to leave. The numbers of the Roget topics which contain the word "point" are the following: 8, 26, 32, 67, 71, 171, 180, 182, 193, 253, 454, 550, 574, 620, 842, 939; and a successful move in this game will be held to be made when a topic number in the "point"-list is also found in the topic-list of another word in one of the sentences. This number will then be said to define the relevant context of "point". Work with Roget will show that in the first sentence the topic 842, "wit", is also found in the topic-list of "facetious"; in the second sentence, the topic 253, "sharpness", is also found in the topic-list of "steeple"; in the third sentence, the topic 620, "intention" is also found in the topic-list of "project"; and in the fourth sentence, the topic 454, "topic", is also found in the topic-list of "argument". So in this case the game successfully works out.

In the "omnis" case the game also successfully worked out. In the range of about ten translations given of "omnis" by this method, the output included "in the main" and "for the most part". And, given the whole context of Book I of Caesar's *Gallic War,* this is exactly what that first use of "omnis" does mean (though they never told us this at school), i.e., "In the main, broad and large, Gaul is divisible into three parts." In another such case, the translation of "incurvo" in Virgil's "Agricola incurvo terram dimovit aratro" came out not as "curved", as the commentators say, but as "bent" or "twisted". But, as stressed above, in many cases the method does not work.

* * *

After such shocks it will no longer seem surprising to learn that the use of computers to process language for mechanical translation, though it has not, as yet, produced workable mechanical translation, has suggested, to the research workers in the field, a number of fundamental new possibilities about the nature of language. Yngve, of M.I.T., for instance, has asserted that at any point in the course of uttering a sentence the minimum number of steps by which a sentence (though not necessarily the sentence intended) could be completed must not exceed a certain very early limit (i.e., of the order of four). Ida Rhodes, of the United States Bureau of Standards, together with the Harvard Computation Laboratory under Oettinger, has shown that quite simple conditional-probability chains can be used to mechanize syntactic analysis; the Lambek-Bar-Hillel calculus illuminates (for me at least) the mechanics of the adjective-noun and noun-verb relations; so does the Parker-Rhodes hypothesis (if true) to the effect that the roles of morphemes in syntactic structures can be represented as abstract sets—and thus as points on a finite product-lattice—which lattice is common to all languages.

Then there is the "language is a thesaurus-web" hypothesis (if true and manageable) exhibited above; and developing this there is the Hesse hypothesis for making a machine find analogies. This says that a mechanizable intersection-algorithm on a thesaurus-type classification-system can be used to give the last term in four-term analogies; e.g., in the analogy, "area" is to "volume" as "square" is to "cube". Then there is the Chomsky hypothesis (if true and manageable) that a full language can be mechanically constructed by deriving it mathematically from a small set of key-forms or "kernels"; and now also there comes a phonetically derived hypothesis by Guberina of the Institute of Phonetics, Zagreb, to the effect that the phonetics of intonational form, when generalized, show that there is a single basic semantic form of human communication, consisting of two pairs making four points of full emphasis, and that this basic pattern can be detected by a machine. And this, as in the case of theory of classification, is only the beginning of the change in our vision of the fundamentals of symbolism.

But the most basic thing which comes out of all this is that the famous division between the two "cultures"—that is, between the artist and the scientist, does not go down beyond a certain point; it is not complete. I have stressed the function of the computer in vision-changing; but *par excellence* it is the artist who changes vision. I have shown, in the experiments, that in every case the point of start for this vision-changing is dreaming up a novel way of defining a "world". But what different thing, what else does the painter do when he establishes his palette, determines his technique and selects his painting instruments, than define a visual world, with which he either does, or does not succeed in changing our vision? And the musician? And the choreographer?

When computers have been programmed to nose out more new mathematical proofs, when one-pack patiences have been mechanically played to simulate the nature of a control system, when a computer is used, but reliably (as it will be), to paint pictures and write poems, and equipped with a machine-sized thesaurus to translate and therefore comparatively to identify differences of context in metaphysical and theological statements in different languages, when all this happens, will the programmer, the analytical wielder of this new mathematical paintbrush, be an artist, or will he be a scientist ?

The Times Literary Supplement—April 27, 1962