A simulation model of human information retrieval *

NICO H. FRIJDA
Amsterdam University, Dept. of Psychology

LAMBERT MEERTENS
Mathematical Centre, Amsterdam (Netherlands)

Le modèle présenté est basé sur l'analyse de protocoles de pensée parlée.
Le modèle consiste en un réseau non ordonné dans lequel la recherche de l'information s'appuie sur les heuristiques dépendant de la tâche.

The present paper is concerned with the simulation of human information storage and retrieval. We will present the outline of a program which is still under development. The program is based primarily upon theoretical conceptions concerning the psychological process. With its help we intend to explore the consequences of these conceptions. The human behavior to be simulated is derived from thinking-aloud protocols. The subjects in our experiments had to solve problems such as: what is a bottle? what is the shape of a mango-fruit? what is the similarity between a crater and a tower?

A model of human information retrieval should, in our opinion, incorporate the following properties. First of all, it should enable the system to produce output statements concerning relationships between elements of information; it should be capable of judging the correctness of such statements, when presented to the system. A traditional associationist, or slightly less traditional neo-behaviorist model is unable to do this. It could explain that "apple" makes me think of a fruit, or, at most, that it makes me think of both the notions of "fruit" and of class-membership, but it could not explain that I know that an apple *is* a (member of the class of) fruit. Searching along linear lines of associations ("apple"—"class membership"—"fruit") can only lead to confusions, as Selz already has cogently shown (cf. De Groot, 1965). An association to both apple and class membership could be Jonathan-apple as well as fruit.

Stored information must consist of somewhat more ordered structures. A way to represent such structures is by an association network in which the elementary data are the nodes, and in which the connections between the nodes represent

* The research reported in this paper has been supported by the Netherlands Organization of Pure Research Z.W.O.

the relationships between those elements. The nature of the connections in the network is, thus, highly differentiated. The functional element of information in this network is not the elementary datum (such as "red" or "fruit") but a complex of the form *A*, *R*, *B*, "*A*" and "*B*" denoting elementary data, and "*R*" any relation or attribute. "*A*, *R*, *B*" may represent "apple, is-a, fruit" or "is-a, is-a, attribute" or "Amsterdam, capital-of, Holland" or "lemon, color, yellow" or "*b*, follows, *a*". It is evident that the nodal elements *A* and *B* may themselves represent highly elaborate complexes of *A*, *R*, *B*-structures. Incidentally, while the existence of an *A*, *R*, *B*-unit implies that *B* can be reached through *A*, the reverse is not necessarily the case; this is a not implausible, but to some extent arbitrary, assumption.

The second requirement for a model of human memory is that it allows of inferential information retrieval. When I cannot remember (or maybe do not even know) what shape your car has, I may plausibly infer it from my knowledge of the common shape of cars. Or, when I do not remember the immediate cause of World War II, I may infer it from my knowledge about events preceding World War II, and their antecedents and consequences. I know and utilize the mutual relationships of causes and antecedents etc. The inferential capabilities of the human information storage and retrieval system are a necessity when modes of access are limited (as they must be), while still allowing a large measure of flexibility of retrieval and of classification of stored information.

Inference of the kind described is possible thanks to two features of the system: interpretativity, and knowledge of logical structures. All concepts in the system are represented as collections of properties which, together, constitute their meaning. This meaning can be visualized either as that part of the network extending from a given node (the concept), or as an attribute-value list headed by some symbol indicating the concept. The definition of a concept by means of its properties (that is, its *R*, *B*-connections) is recursive. Each property, that is to say each relationship or attribute, and each associated concept, is defined in a similar way by sets of properties, until a final level is reached. This final level is supposed to consist of sense-data, elementary feelings, and elementary routines. At any time, a concept can be replaced by, and gives access to, some or all aspects of its meaning. Concepts consist of sets of criteria for acceptance of data as specimens of the concept. Conversely, further information concerning a concept can be retrieved by way of its properties. When asked about the edibility of a mango, which I do not know, I can make a guess since a mango is a fruit and fruits usually are edible.

Not only the elements of the system can be thus interpreted; so can the relations or attributes. As a matter of fact, the system does not know any essential difference between elements and attributes. A given concept functions as an attribute when it occurs in the middle position of an *A, R, B*-complex, or when one of its properties is "is-a, attribute". However, some relations have aspects of meaning which are rather involved, and which permit extensive inferential reasoning. These aspects are logical or quasi-logical structures. The system may know "apple, is-a, fruit"; it also knows what "is-a" means. It can, and often will, retrieve this meaning: an entire bundle of rules, stating, for instance, that any property defining a class (i.e., the right-hand member of any "*A*, is-a, *B*"-structure) necessarily pertains to any element or subclass of that class (i.e., the left-hand member). These rules are the bridges which permit utilization of fruit-knowledge when thinking about mangoes.

Thirdly and finally, any information retrieval system which uses inference must manifest selectivity in memory search. It must be able to select those associates as candidates for inferential assistance which show some promise to be useful. In other words, the system must possess memory retrieval strategies. At any rate, the psychological model must manifest those memory retrieval strategies apparent in thinking-aloud protocols.

The strategies human beings use when solving problems of this sort are quite task-specific. They are dictated by the properties of the questions asked and the task field they are to operate in. It would be unfeasible as well as implausible to preprogram these strategies for all task types the memory system might be confronted with. It is also implausible that human beings have learned specific strategies for all different sort of questions they may have been asked.

Nor need they do so. The task-specificity of memory retrieval strategies (and, to a large extent, problem solving strategies in general, cf. Elshout and Frijda) indicates that the task itself—the meaning of the question—implies and dictates the strategy to take. Or, rather, the strategies follow from the meaning of the question and the structure of the stored information. Strategies are *constructed*, for the most part, *ad hoc*, when and mostly while solving the problem. The principle of their construction is that the task selects the possibility relevant aspects of the logical structures, whereupon these aspects dictate the heuristic means and problem transformations to be tried.

To illustrate: a request for a definition implies (at least in the case of concrete concepts) a request for properties pertaining to all members of a class; one has to suppose this information to be stored together with the concept "definition". A request for "property pertaining to all members of a class" matches with an aspect of the concept of class-membership. Consequently, search for examplars of the concept whose definition is requested, and subsequent search for the common properties of all, is a plausible strategy, found by this, in principle, simple means.

Within the memory structure just described, search is effected by means of a compound "stimulus" which, if all goes well in an easy way, matches one or more compound memory elements. The compound stimulus consists of an incomplete *A*, *R*, *B*-structure (such as "apple, color,?"), or similar but more elaborate structures (e.g. ((mate, *R*, ship), (pilot, *R*, ?))). This stimulus serves as the trigger for memory search, as well as the basis for testing *post hoc* the success of some roundabout search: is the information found indeed the kind of thing I was looking for. It is evident that the testing function could not be fulfilled by a simple stimulus in the associationist sense, nor by a collection of such stimuli. The structure in the stimulus is essential for this purpose. As a matter of fact, this compound incomplete stimulus is precisely the same thing as Selz's "anticipatory schema" (cf. De Groot, 1965).

Direct search in this way is rather trivial, and success is an exception. For retrieval in case the direct way fails, one simple strategy is used by the system. Quite generally, the system endeavours to find some content of which the desired information is an implication. More precisely, the system searches for some information which is connected to the information requested by a relationship which is equivalent to an implicatory relationship. If asked for the function of a mango, the information that mangoes are fruits is relevant, because class-membership hides an implication of the correct kind, through its aspect "any property of a class is a property of any member or subclass of that class". In the present form of the program, this search for implications is performed somewhat rigidly. Only the correctness of the direction of the implication and the correspondence of its terms with those of the problem, are investigated, but the fact of an implicatory relationship is signalled as such. A more flexible and more naturalistic method can fairly easily be implemented, however. Also in preparation is an other mode of operation, in which "*A* implies *B*", where *A* is the desired information, can be used also. This will lead to a probabilistic mode of operation, in which more or less likely information will be retrieved, and tested for appropriateness at a later moment.

This search for information which implies the information requested is one of the essential ingredients of our model of the human information retrieval process. There is a second essential process feature which is, at present, only very incompletely incorporated. In human memory retrieval, when searching for an item of information, plausible answers are tested as to their fittingness with respect to what is asked. When requested to name a great contemporary Greek poet, I might wonder whether the composer Theodorakis might fit the description. That is, the system generates its stored meanings of "great contemporary Greek poet" and "composer Theodorakis", and investigates whether the overlap of meanings is acceptable,—criteria for acceptability varying according to a number of circumstances. This principle of fittingness again permits a large measure of flexibility of processing, and an economical manner of working from beginning and end in solving subproblems, which we hope to include in our program.

To illustrate the operation of the program, as it stands now, let us follow the processing of a simple question in a simplified memory space. We present the output of the program, with comments.

retrieve (cactus, color, ?) when the contents of memory are:

(cactus, is-a, plant)
(color, is-a, property)
(color, synonymous, hue)
(synonymous, aspect, ((*A*, *R*, *B*) provided (*R*, synonymous, *S*), (*A*, *S*, *B*)))
(is-a, aspect, ((*M*, *P*, *V*) provided (*M*, is-a, *C*), (*P*, is-a, property), (*C*, *P*, *V*)))
(plant, color, green)
(is-a, inverse, exemplar)
(inverse, aspect, ((*A*, *R*, *B*) provided *R*, inverse, *S*), (*B*, *S*, *A*)))
(synonymous, inverse, synonymous)
(inverse, inverse, inverse)

The maximal number of transformations that may subsequently be applied to a problem has to be specified. In this case we choose it to be equal to one.

Because the problem (cactus, color, ?) is not compound, the program immediately starts a search for elements of information that answer to the description (cactus, color, ?) or imply a structure answering to it. In doing the latter, it starts with investigating the associations of the concept, "cactus", and after that those of the attribute, "color".

(cactus, color, ?)

interpret cactus
(cactus, is-a, plant)
interpret is-a
(*M*, *P*, *V*) provided (*M*, is-a, *C*), (*P*, is-a, property), (*C*, *P*, *V*)
test for implication: cactus is filled in for *M*, color for *P*, and the possible values of *V* are asked for.
(cactus, is-a, *C*), (color, is-a, property), (*C*, color, *V*)
This is a new problem into which the original one has been transformed. The first thing to do is to test whether (color, is-a, property) holds. As this is the case, the program tries to tackle next the most promising atomic problem:

(cactus, is-a, ?)

interpret cactus

(cactus, is-a, plant)
This answers to the description and yields the answer "plant". As the search is rather exhaustive, the program will go on searching for other solutions.
interpret is-a
(*M*, *P*, *V*) provided (*M*, is-a, *C*), (*P*, is-a, property), (*C*, *P*, *V*)
Because no further transformations are allowed by the depth-of-transformation limit, the test for implication is suppressed.
(is-a, inverse, exemplar)
interpret inverse
(inverse, inverse, inverse)
(*A*, *R*, *B*) provided (*R*, inverse, *S*), (*B*, *S*, *A*)
All this leads to nothing.
result: plant
The next thing to be tackled is:
(plant, color, ?)
This yields: green (account of the searching process omitted here).
So we find: green
The search continues, but no more values will be found.
(is = a, inverse, exemplar)
interpret color
(color, is-a, property)
(color, synonymous, hue)
interpret synonymous
(*A*, *R*, *B*) provided (*R*, synonymous, *S*), (*A*, *S*, *B*)
test for implication:
(color, synonymous, *S*), (cactus, *S*, *B*)
(color, synonymous, ?)
This yields: hue
(cactus, hue, ?)
This happens to yield nothing at all.
The final result, (cactus, color, green) is added to the memory store.

It deserves mention that the program may, under circumstances, solve problems during a second run on the same problem, when it failed on the first run, because partial results have been added to the store. In the above case, the problem would have been solved directly had it been presented the second time.

The present program is, of course, a close relative of other inferential programs, which usually go under the name of natural language programs, or semantic programs (e.g. Lindsay, 1961, Raphael, Quillian). It differs from those, either because its intent is psychological simulation rather than understanding natural language, or because of its feature of interpretativity, or both. Moreover, we think the fact that the specific strategies are derived from the data structure is present in none of these other programs.

REFERENCES

De Groot, A.D. (1965), *Thought and Choice in Chess*, Mouton, Den Haag.

Elshout, J.J. and Frijda, N.H. (1966), The origin of problem solving methods, *Proceedings, 20th International Congress of Psychology*, Moscow, vol. Heuristic Processes in Thinking.

Lindsay, R.K., *Towards machines which comprehend*, U. of Texas, report.

Raphael, B. (1964), *SIR: a computer program; for semantic information retrieval*, Thesis, M.I.T., MAC-report TR-2.

Quillian, R. (1966), *Semantic Memory*, Bolt, Beranek and Newman, report AFCRL-66-189.