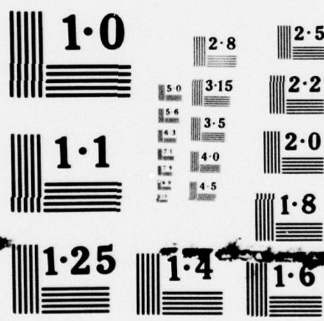
AD-A058 904 MARYLAND UNIV COLLEGE PARK DEPT OF COMPUTER SCIENCE F/G 5/10
THE IMPORTANCE OF MULTIPLE CHOICE.(U)
JUN 78 C RIEGER N00014-76-C-0477
UNCLASSIFIED TR-656 NL

1 OF 1
ADA
058904

END
DATE
FILMED
11-78
DDC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

LEVEL II

②

COMPUTER SCIENCE
TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

(2)

(14) TR-656

(11) June 1978

(9) Technical rept.

(6) The Importance of Multiple Choice.

(10) Chuck Rieger
Computer Science Department
University of Maryland
College Park, Maryland 20742

(12) 24p.

(15) N00014-76-C-0477,
NSG-7253

A paper presented at the TINLAP-2 meeting, Urbana, Illinois, July 1978.

The research described in this report was funded in part by the Office of Naval Research under grant number N00014-76C-0477, and in part by the National Aeronautics and Space Administration under grant number NSG-7253

409 022

The Importance of Multiple Choice

Chuck Rieger
Department of Computer Science
University of Maryland
College Park, MD 20742

Abstract A level of organization of inferences in which competing plausible alternatives can be compared, and all but one actively rejected, is a very important aspect of any comprehension model. Multiple choice inference structures help the model stay tuned to the comprehension context, and help establish a framework in which inference producers are less likely to outstrip inference consumers, a common problem of inference systems. I look at four different modelling areas in which the same issues dominate, and suggest that the structure of knowledge is greatly influenced if one adopts the point of view which places emphasis on multiple choice.

1. Introduction

In any comprehension model, there must be a more or less equal number of producers and consumers. That is, for every inference the system is capable of generating, there must be some other part of the model that is capable of responding to the new inference. Still another way of putting it is that any comprehension system must be capable of understanding its own inferences.

This paper is about problems of inference consumerism. I, and I suspect most others, always find it easy to write generative rules of inference (ones that trigger on original inputs to the system), but extremely difficult to deal with all the loose ends created in the process. Most of the inferences are really insightful for the first level or two, but then it becomes very tedious to keep the system closed; most inferences that get generated simply fall between the cracks. This is not pleasing as a model of comprehension.

The problem usually seems to be that we are tempted to write overly schematic inference patterns, ones that can respond to infinitely many other inferences besides the small, finite set we have in mind when writing the rule. The infinity usually comes from too liberal a use of pattern variables in the inference schemata. In the system that results, we get "comprehension" as we knew we would on the finite world which motivated the development of the inference schemata. But if the domain is shifted only slightly, while the producers still produce, there are no consumers for them! Doug Lenat discovered this problem in AM, where he had constructed a nice set of inferences (heuristics, in his case) out to several levels, but where, after the range of his forethought had been exceeded, the rule system's open-ended nature reared its head; at that point, the producers had outstripped the consumers.

The same type of problem has always bothered me. It takes only one afternoon to write a collection of rules to "comprehend" a passage from, say, a children's story. But what does one have after finishing? It's hard to say, because after one or two small domain shifts, or after the system runs too long, the consumers become mismatched to the producers. This is not the way humans reason; humans somehow keep their inference systems closed so that producer output somehow always wraps around and manages to tickle consumer patterns, no matter how long the system runs.

In this paper, I want to show how inference producer-comsumer problems might be eased by multiple choice inference structures. This is the main thread. Another point I want to make is that active rejection of all alternatives but one, performed by multiple choice inference structures, is perhaps the most important and theoretically interesting topic of both comprehension and learning. I try to convince you that the development of the typical production rule-, script-, or frame-based style of encoding knowledge is only the first step in a comprehension system. The second step is to build additional inference structures that superimpose multiple choice and active rejection.

I do not fully see the connection between multiple choice and the inference producer-consumer problem. But I would be happy to stir up more discussion of the whole topic of decision structures in comprehension models that must deal frequently with alternatives. My purpose is to call attention to this important second aspect of comprehension, since I feel it has been largely overlooked in the rush to represent base knowledge using production rules, frames or scripts.

I discuss the same issues from four different points of view, in four modelling areas in which inference is important. If, in fact, multiple choice and active rejection as illustrated in these example domains are important issues, there are implications for the structure of large rule, frame or script bases. In the last section, I try to summarize these implications.

2. Example

Suppose I hear "John pounded Mary on the back." What do I do that can be characterized as "comprehension", and, more important, when and how do I know that I've gotten the point. What I do seems fairly obvious: I somehow expose this input to a system of inferences about one person pounding another, probably conditioned by prior context, and monitor the results, looking for "hits" or connections with other patterns previously generated or predicted.

Presumably, my personal system of inferences has been build gradually, more or less one small chunk of inferential knowledge at a time, for a period of many years. When I was 3, perhaps the only cause-effect explanation of an "X pounds Y on the back" event was that Y was choking and X was trying to help. Perhaps when I was 5, I saw that this could be an effective means of physical retaliation, when the other guy wasn't looking! Perhaps later on, I discovered it was also an expression of congratulations, and still later that it was also part of the ceremony of certain religious sects.

Suppose, then, that these four explanations of pounding have been tucked away in a large system of rules. In some loose sense, I will always be able to comprehend instances of pounding, because I can always proffer a plausible causal explanation, whether or not it is an appropriate one. But there is an important difference between offering a plausible explanation and "comprehension". In my view, "comprehension" refers more properly to an activity in which the experiencer convinces himself that, among all the possible interpretations, he has found the most appropriate one in the context at hand. In other words, whether or not he is in fact correct in his interpretation of an event, a human can usually feel the "click" of comprehension that tells him he has locked onto the right interpretation.

Comprehension, therefore, seems to be more a process of discerning among plausible alternatives at each step, rather than simply generating them and following one arbitrarily. This means that, if we have four plausible explanations of pounding, we can only comprehend instances of pounding after having taken all four into account and actively rejecting three and adopting one. But to do this, we must have some basis for comparing the four plausible alternatives; they cannot be truly separate inference rules in a large rule base, even if they entered one at a time historically. They must rather have some mutual structures in common that allow them to be contrasted along various dimensions.

If this line of reasoning is correct, the conclusion seems to be that comprehension cannot occur until all plausible alternatives have been considered, compared in context, and all but one actively rejected, for good reasons. If this is in fact the mechanism of comprehension, then the evolution of a rule system must go as follows: a new rule is acquired, used several times somewhat at random to see if it really works; if the rule seems correct, it is woven gradually into some sort of discrimination structure in which it represents a new alternative explanation for an event for which several other explanations already exist. It is in this significant process that diagnostics - questions that can differentiate all the various plausible explanations - become compiled. (I also believe that "learning" is most accurately characterized by this process of weaving a new plausible explanation of an event into a discrimination structure that can actively reject all but one explanation for any instance of the event.)

I believe that it is such discrimination structures that distinguish a child's comprehension from an adult's. A child is largely a collection of unorganized rules, each of which is a plausible explanation of some event. A child can therefore offer plausible, but perhaps inappropriate explanations of many events. But, lacking parts of the data structures and mechanisms of active rejection of all but one competing plausible explanation, the child perhaps does not comprehend in the same sense an adult does; he is not as confident in saying whether or why his explanation is the appropriate one in context. (Alternatively, perhaps his active rejection machinery is working, but his limited repertoire of plausible alternatives for any given event causes miscomprehension. In either case, the active rejection mechanisms play a central role.)

The point, therefore, is that when I hear something like "John pounded Mary on the back," I do not simply ask "Why did he do that?" Instead, I ask, "Why did he do that? Pick one: (1) to remedy Mary's choking, (2) to vent anger at Mary, (3) to congratulate Mary for some accomplishment, or (4) to save Mary's soul." This will force the differentiation and active rejection mechanisms into play, and, as a method of writing inferences, will force us into identifying the consumers of the inference (i.e., if the answer is (1), what do we do next?).

Multiple choice inference systems are, I suspect, another facet of the problem addressed by Minsky's frames and Schank and Abelson's scripts and plans. The tenet of these theories is that comprehension is mostly a matter of knowing what the ingredients of a situation or event are so that a system knows what details are important to fill in or react to. Frames and scripts are both attempts to focus inference along the paths deemed most relevant by the current context, a.k.a., collection of active frames or scripts.

Demanding that frames and scripts not only define what aspects of a situation are relevant, but also restrict the range of possible outcomes to an explict set of alternatives for which consumers are guaranteed to exist is critical to a comprehension model. For, in even a frame or script-based model, although there is tight control over focus of attention, without multiple choices and the accompanying differentiating processes, there is no guarantee that the adopted line of reasoning will be the most appropriate one. In other words, I think there are two issues: how to organize knowledge as production rule-, frame- or script-like structures to define relevance and focus in an inference system, and (2) how to superimpose the multiple-choice, active rejection mechanism on top of this knowledge. Frames and scripts relate more to base representation (that used to express the individual rules of a domain); multiple choice and active rejection relate more to the mechanisms that manipulate the base representation; the two theories are complementary, but somewhat independent. (I suspect that similar feelings motivate Davis, for one, to propose meta-rule structures. However, I would not classify my multiple choice and active rejection hypotheses as meta-knowledge.)

3. Some Examples

Illustrations of the importance of multiple choice techniques in comprehension can be drawn from almost any domain. I want to consider them in four model categories: meaning-based parsing, problem reduction problem solving, inference, and interactive systems. Most of the discussion relates to projects with which my group has been involved.

3.1. Multiple Choice in Meaning-Based Parsing

The traditional approach to parser construction, whether at a syntactic or semantic level, has been to write a system of rules about linguistic and world knowledge. The rules, which describe sentence and concept-level constructions, are then interpreted by a more or less uniform interpreter. Most such systems can be thought of as production systems, with various control twists superimposed. In writing the rules for such a system, the experimenter is encouraged to seek and express the regularities of meaning and syntax; he is implicitly concerned with the conciseness and generality of the rule system.

I believe that the entire concept of sentence-level rules, interpreted by a uniform interpreter, is ill-fated as a model of human language understanding. It is my belief that most of the richness of both the content of language, and of the control paradigms which interpret it, derive more from word-level knowledge than from sentence-level knowledge. As a result of this belief, I believe that a natural language parser should be built as a population of autonomous word experts, each of which is capable of discriminating its intended sense in context.

In the Sense Expert Parser we have built, there is essentially one expert process for each word of the vocabulary. The parse of a sentence in this framework amounts to calling up the experts which the input sentence references, informing each of its location in the sentence (i.e., telling it who its neighbors are), then turning the population loose as a collection of asynchronous, parallel computations which may interact both among themselves and with several levels of model context (the story comprehension model, in our case). Being autonomous, each word expert is free to grow to any size to accommodate all the nuances and idiosynchracies of its word's usage, as well as to accommodate the possibly numerous "standard" senses of the word. Since each expert can be paged from an essentially infinite disk file, and since only as many experts as there are words in the input sentence need be in memory simultaneously, the theory is also quite practical. We have written several reports describing the theory and implementation of the Word Parser [], so I will not go into any of the interesting problems of parser control here.

What I do want to emphasize, however, is that this view of language leads naturally to word expert structures that do a lot of multiple-choice question asking, to make inferences that narrow down the possible sense of each word in context. In fact, although we are redesigning the system now to look more like a collection of CONNIVER coroutines, the internal structure of each sense expert always has been that of a discrimination network whose terminal nodes are meaning fragments that come to be stitched into the sentence-level meaning translation produced by a parse. At each node of the discrimination net, there is a multiple choice question about either the local sentence environment or the model environment. A typical question of one expert to another is: "What is the likely semantic category of the referent of your sentence component?: PHYSOBJ, HUMAN, ABSTRACT-CONCEPT. Based on the outcome (specifically, the active rejection of two of these three alternatives), the asking expert poses another question, and so on, until he is confident that enough of his context has been probed to make an accurate inference about the intended *meaning of the word.*

I feel that active rejection is fundamental to language understanding. At each step of sense disambiguation, there must be clear evidence for the superiority of one interpretation over all others, and hence an accompanying query that asks not absolute questions, but relative ones capable of differentiating the alternate interpretations.

It is also my belief that the synthesis of individual word senses, acquired one at a time through more or less haphazard experience, into these multiple choice sense expert structures is the most basic mechanism of language acquisition and learning.

3.2. Multiple Choice in Problem Solving

In open-ended domains, e.g., everyday life for a human, where there are not always obvious solutions to problems, multiple choice structures play a role analogous to that in parsing. Suppose a problem solving system is confronted with a task: "Move object X from its present location Y to a new location Z." If we are dealing with blocks on a table top, the problem has an obvious solution. If however, I approach you with this task as a real world problem, if you decide to solve it, you will undoubtedly ask some questions. The point of the question asking will be to find out more about X, Y and Z. Why? Undoubtedly to enable you to select an appropriate strategy, so that you can see several steps ahead, say, to give me an estimate of how much it will cost. Specifically, you will probably ask such questions as: "What is X?", "Where is it now?", "Where is it to be moved?", you will have to select from among the hundreds of strategies you know for moving objects in the world. This is an obvious problem, but it is a requirement of an open-domain problem solver that is not ordinarily experienced in a restricted domain problem solver, such as one for the blocks world, or one for electronic circuit analysis.

My point is that building an open-domain problem solver leads to a different set of issues, and hence a different underlying organization from building a restricted domain problem solver. Just as word senses, and active rejection of inappropriate ones becomes the central issue of an open-domain parser (as opposed to, say, a blocks-world parser, where every word has a relatively unambiguous interpretation), strategy selection (that is, active rejection of all but the most contextually appropriate strategy) become a main topic in an open-domain problem solver. Just as with word senses, there must be not only the strategies themselves, but also structures for differentiating ones with similar goals according to their appropriateness in specific situations.

I believe that it is the same sort of evolutionary synthesis of isolated strategies into multiple choice strategy selection structures that reflects learning in problem solving. There is one critical difference, however, between selecting appropriate strategies via multiple choice differentiators and selecting appropriate interpretations for words during comprehension. It is that in problem solving, while optimality of the selection process is highly desirable, it is not critical to the ultimate success of the endeavor. (I could, after all, move a mountain with a teaspoon.) In comprehension, on the other hand, optimality is everything; the speaker or writer is in the business of transmitting enough information to allow the correct (optimal) selection of each word's meaning and of each thought's interpretation. If you miss the "optimal" interpretation, you do not comprehend, at least in the way the speaker or writer intended.

3.3. Multiple Choice in Inference

As I pointed out in the example at the beginning, making inferences is not just a matter of finding plausible explanations, i.e., of simply firing off production rules, or applying frames or scripts whenever their preconditions match the situation. That, certainly, must be the first step. But the important difference between this step and "comprehension" is that comprehension demands a next step wherein differentiating questions are posed to decide upon the most relevant of the plausible explanations. If you are writing scripts, and several of them are triggered by "X got on the subway" as plausible explanations, the next, and by far most important step, is to invoke some sort of decision procedure to decide which to adopt at the controlling script. If you are writing production rules for infectious diseases, and several rules fire on "X's temperature is above 102", the next step should be to examine the competing plausible explanations differentially, via a decision procedure aimed at the relative disambiguation of the explanations. Cumulatively, the sequence of local, relative disambiguations leads to a global, more absolute interpretation.

There are many levels of a comprehension model at which multiple choice inferences are essential. Fauser has made some interesting points about text relevance, and has developed a type of prediction mechanism he calls "interest patterns". Roughly speaking, for each new input to the system, in addition to the story-level prediction/fulfillment mechanism that is running (i.e., the level that from "Minnie went to the cupboard" predicts that "Minnie might be looking for food"), a relevance-level mechanism is constantly generating interest patterns. These are patterns that characterize the nature of information the comprehender would expect to hear more about, rather than the information itself.

Fauser's favorite example, "John shot Bill. He died", illustrates the importance of interest patterns in certain reference tasks. Here, knowing (or predicting) whether Bill died is a story-level task that the system might well want to perform. But, at the relevance level, it suffices to note that Bill's physical state is at issue, regardless of the outcome. When the "he" in the following sentence refers to a person whose physical state is being characterized, the interest pattern makes it reasonably clear that information about a person's physical state in the context of the first sentence most properly belongs with the referent of "Bill".

I mention interest patterns and relevance-level expectations because, for these types of predictions to function properly, there must be an element of multiple choice. The system cannot simply post an interest pattern saying "I am now interested in Bill's physical state", and expect that pattern to be useful in all possible future contexts. (Suppose the next sentence were "His trigger finger ached", a statement about someone's physical state, but probably not Bill's in that case!) Rather, it must say: "I am interested in Bill's physical state. Specifically, it is now relevant to know whether Bill was (a) totally unharmed, (b) mildly injured, (c) seriously injured, or (d) killed. I will be satisfied only when a statement of physical state that relates to one of these characterizations is discovered." These choices can be quite valuable to the system, since their definitions can be compared to compute differences. The differences can then serve as sources of diagnostic questions that are actively posted to monitor future inputs.

Multiple choice inferences seem especially well-suited to the task of tracking down causal explanations. Suppose the input is "Minnie was hungry." There can be many reasons, each represented by a cause-effect pattern somewhere in the comprehender, why a person might be hungry. The import of this input can vary tremendously according to which explanation is chosen. For example, Minnie might be an overeater, always hungry; she might be hungry simply because it is dinner time, and she hasn't eaten since lunch; or, she might be chronically hungry because she has too little food. To "comprehend" this sentence therefore seems to demand that the "X is hungry" inference pop up and actively pose questions that could help differentiate these three conditions. It might, for example, ask about Minnie's financial well being; if the answer is that Minnie is pretty well off, then the third plausible explanation can be actively rejected. It might then ask whether or not Minnie is known to be a glutton, how long it's been since Minnie's last meal, and so forth. The point, again, is that in order to make a good inference, the system has to be aware of all the plausible alternatives, and has to select among them by invoking diagnostic questions before making the inference. In doing so, it will also help explicitly identify the appropriate consumer.

3.4. Multiple Choice in Interactive Systems

The case for multiple choice data structures in interactive CAI or CAD systems, those which deal directly with some external intelligence, is perhaps the most obvious and familiar. In a typical exchange, the system askes the user or designer a question. Based on the response, a model construct is generated, or some sort of model response emitted.

Restricting inputs to such a system to precisely those forms for which there are well-defined consumers is paramount. For example, suppose we are in the middle of a digital design, and the system needs to know the frequency at which a certain counter must count. If it asks the user "At what frequency?", and the user responds "pretty fast", what should happen? Alternatively, if the user says "I don't know exactly, but somewhere between 10 mhz and 100 mhz", what should the system do? In the first case, the interaction has probably been fruitless, because, even though there is some small amount of information content to the response "Pretty fast", it will probably not reference a meaningful concept (or even meaningful symbols) in the model. In the second case, although mhz units may be understandable to the model, there may be a critical difference in the type of circuit required to accommodate a 10 mhz versus a 100 mhz signal. Since questions in a CAI or CAD system are invariably posed for specific reasons, the answerer must be made aware of those reasons so that he can provide a consumable answer. The most direct way to ensure this is to have the asker force multiple choices on the answerer.

My point, again, is that multiple choices are necessary both to guarantee that information producers can only produce information that is consumable by some consumer in the system, and to define diagnostics needed to distinguish the important cases. In parsing and inference, this means distinguishing the most probable plausible inference or word sense from all rest; in problem solving it means distinguishing the most appropriate strategist to solve a subgoal; and in a CAD or CAI system, it means forcing the answerer into providing an answer that will allow the system to select a well-defined next action.

4. Implications

I suspect I've beaten the issue long enough. What are the implications?

First the hypothesis, quickly restated: Good comprehension means ensuring that a system's information producers and consumers form a closed system, and that, whenever confronted with more than one plausible alternative, there be a mechanism for actively rejecting all but one on the basis of multiple choice diagnostic questions. To me, this suggests that a production system-like approach to writing a rule base, in which there are likely to be many instances of competing rules that are kept insulated from one another, is not adequate for most types of comprehension. Instead, there need to be mechanisms for active rejection of all but one of the first-phase competitors. Such a mechanism can be built into the system in one of two ways. One way is to invoke arbiters on the fly, when competing interpretations surface. An arbiter, designed for just such a case, would emit a sequence of local diagnostics, and hopefully actively reject all but one interpretation and, in the process, identify a consumer of the interpretation.

A second realization of multiple choice, and I think a better one, is to structure the rule base in a way that inherently defines arbiters. In such an organization, rather than maintaining numerous rules with similar patterns (rules which would compete in many situations), all possibly competitive rules are organized into a discrimination structure which implicitly knows of all the competitors, and most importantly, knows good contextual diagnostics for actively rejecting all but one in any given context. In this style of rule base, the large collection of independent rules is replaced with a somewhat smaller collection of rule clusters, each cluster being represented by an expert diagnostician. Rather than the reflexive firing of competing rules, there is instead the reflexive firing of diagnosticians that first select which rules to fire, then fire them.

Generally, the point I want to make is that it is the evolution of this second level that transforms a system from a gangling data base of plausible inferences into a system capable of locking onto good paths by differential comparison of related alternatives. Before the transformation, the system reacts; after the transformation, it comprehends, since it can not only arrive at plausible explanations, but it can also locate, compare, and actively reject all the competing explanations.

What this means to the designer of rule bases using production rules, frames, or scripts is that the design is a two-step process. First, he engineers the knowledge into one of these representations, not paying particular attention to similarities among subsets of rules. He then examines or runs the system, noting the subsets of base knowledge that fire for test cases representative of the types of situations the system will experience. He then builds a decision structure whose queries will probe context in the most incisive ways possible to discriminate the most appropriate alternative. As the base knowledge evolves, so must the discrimination structures. (Even in the absence of base knowledge evolution, evolution of the discrimination level can have profound effects on the efficiency and accuracy of the system.)

The conclusion I would have you draw is that multiple choice inference structures are a fertile research area, and one that has been unduly overshadowed by the recent interest in he structure of individual inference patterns. A well-timed, well-chosen diagnostic question at a multiple choice step can save a lot of elaborate control structure and wasted energy spent backing up, or looking for an appropriate consumer.

Acknowledgements

I thank the members of the CSA research group for their interest and comments. I also thank NASA and ONR for their support of this and related research under grants NSG-7253 and N00014-76C-0477, respectively.

5. Readings

Davis, R., Interactive Transfer of Expertise: Acquisition of New Inference Rules, in Proc. IJCAI-77, Boston, 1977

Fauser, A., Pronoun Interpretation in a Story Comprehension Model, UOM Computer Science TR-636, 1978

Feigenbaum, E., The Simulation of Verbal Learning Behavior, in Feigenbaum and Feldman (eds.), Computers and Thought, McGraw-Hill, 1963

Lenat, D., AM: An Artificial Intelligence Approach to Discovery in Mathematics as Heuristic Search, Stanford AIM-286, 1976

Minsky, M., A Framework for Representing Knowledge, in Winston (ed.), The Psychology of Computer Vision, McGraw-Hill, 1975

Rieger, C., Viewing Parsing as Word Sense Discrimination, in Dingwall (ed.), A Survey of Linguistic Science, Academic Press, 1978

Rieger, C., GRIND-1: First Report on the Magic Grinder Story Comprehension Project, Discourse Processes, 1(3), 1978

Rieger, C., Five Aspects of a Full-Scale Story Comprehension Model, Findler (ed.), Associative Networks: The Representation and Use of Knowledge in Computers, Academic Press, forthcoming Academic Press, 1978

Schank, R., and Abelson, R., Scripts, Plans, Goals and Understanding, Lawrence Earlbaum Press, 1977

Small, S., Conceptual Parsing for Story Comprehension, UOM Computer Science TR-663, 1978