THE CONCEPT OF COMPREHENSION: FROM SEMANTICS TO SOFTWARE[1]

Michael Scriven
University of California at Berkeley

INTRODUCTION

An argument is given that simple considerations of efficiency for an information acquisition-storage-retrieval-reaction device (IPD) interacting with a complex informationally rich environment will necessitate strategies that ensure meeting the criteria for comprehension. The first section analyzes "comprehension," the second deals with efficiency in IPDs, and in the third, implications of a comprehension theorem are briefly discussed.

Comprehension

The term "comprehension" appears to be essentially synonymous with "understanding" over its main range (although its range is desirably narrower, excluding the use of the latter in "*an* understanding" that is roughly equivalent to "an arrangement or agreement"). These concepts are intimately related to the concepts of explanation and knowledge, so the present result has some significance for explanation-theory and for epistemology. An explanation is a device or process for communicating comprehension and knowledge is the generic state under which fall the various species of comprehension, *inter alia*.

We can identify from previous discussions (Scriven, 1964) a positive and a negative way of precisifying the concept of comprehension. We may give (paradigm) *examples* of it or (paradigm) *contrasts* to it. Five 'types' of comprehension can be exemplified of which it is sometimes said that one or another of them is the 'real,' 'best,' 'fundamental,' or 'only scientific' sense. The present treatment analyzes all five as special cases of one concept, not as different senses or concepts. The same applies to the four main contrasts. (The phrasing is in terms of "understanding'' because of its greater range of correlatives in English.)

Example 1. Understanding an event, phenomenon, or process, such as transistor deterioration or an error in reasoning.

---

[1] This work has been supported in part by an NSF Grant.

Example 2. Understanding a theory, such as information theory.

Example 3. Understanding a natural language, such as English.

Example 4. Understanding an experience, such as the experience of divorce or childbirth.

Example 5. Understanding an entity or a class of entities, such as divorcees, children, sheep, or Chrysler hemi-V-8's.

Contrast 1. Understanding versus ignorance. This is the 'primary response' contrast for most native speakers.

Contrast 2. Understanding versus 'mere knowing,' as in knowing the time or knowing the age of the universe.

Contrast 3. Understanding versus misunderstanding, as in misunderstanding a flow chart.

Contrast 4. Understanding versus believing (or feeling) that one understands, a distinction commonly illustrated in the written answers to course examinations.

An analysis of "understanding" that fits these constraints is essential for any semantic program for outputting (or a disambiguation matrix for inputting) the term but it will not necessarily consist in a definition in the classical sense of a set of conditions that are necessary and jointly sufficient, this usually being impossible, other than trivially, with the important concepts of logic or science such as the so-called ''open concepts'' or ''open-textured'' concepts.

Various oversimplified analyses have been suggested which are obviously inconsistent with the above paradigms. For example, opponents of the scientific or logical legitimacy of the notion of understanding frequently attack it as if it were identical to felt understanding thereby violating Contrast 4.

Much better analyses center on such notions as ''reduction to the familiar," "modeling," "semantic programming," "deduction from laws," etc. However, they all suffer from vagueness or from specific errors which have been extensively discussed in the literature. To avoid vagueness, some attempt at a more operational or behavioristic account is warranted.

Concentrating on Example 2 for a moment, we can ask how it is that we test comprehension or understanding of a theory? We ask the subject questions about it**,** questions of a particular kind. They must not merely request recovery of information that has been explicitly presented (that would test 'mere' knowledge, which must he excluded by contrast 2). They must instead test the capacity to answer 'new' questions. I shall suggest three criteria, not all always appropriate. The alpha criterion involves at least the following capacities which are not essentially distinct: (a) the capacity to make and recognize (what are currently held to be) obvious logical inferences from and translations of stored data (b) the capacity to extrapolate or interpolate any general principles learned; and (c) the capacity to apply the theory to new problem cases to which it is relevant. A related test of comprehension, the beta criterion, seeks signs of organization of the material acquired, especially any sign of novel organization; we sometimes find a student's 'grasp' of the theory most impressively demonstrated by a novel axiomatizing or a new identification of fundemental concepts. The stress here, then, is on the capacity to *handle* novel problem situations and on the capacity to *produce* novel output (where 'novel' is subject-relative). But there is sometimes an emphasis on the gamma criterion: a good perspective or appraisal or evaluation of the theory, its components and its achievements, i.e., (roughly) its relations to other theories and the evidence.

Does this kind of analysis transfer to the other paradigms? There is a very close match with the requirements given by Chomsky and others for understanding a language (Example 3).

Speaking precisely, one might be said to understand French even if one could not speak it. The proof would involve demonstrating that *novel* (French) input produced the appropriate non-verbal or a non-French reactions; in particular, it would require the recognition of synonymous or equivalent passages as such, of true/false assertions and of inappropriate language as such (‘ungrammatical’ in the Chomskian sense). This is close to the alpha criterion. (We can say that the alpha criterion is the appropriate *reaction* criterion, the beta criterion, refers to appropriate *reprocessing*, the gamma criterion to *meta-description* or data-comparison.) The apparently aberrant paradigms 4 and 5 present no difficulty if it is recognized that the notion of appropriate reaction (etc.) is not restricted to verbal performance. This is already clear in the case of understanding a language where obeying a command is the best proof of comprehension. We may say that a mechanic understands a certain family of engines, like the Chrysler hemi’s, without any confidence in his capacity to tell us the theory on which the design is based, solely on the basis of his ability to fix them, improve them, and diagnose them. Similarly, a nurse need know nothing of academic psychology to understand children. She need only be good at ‘‘instinctive’’ diagnosis and treatment. Although the alpha criterion is the most important in these semi-linguistic areas, there is still a clear analogy to the beta and gamma criteria.

Understanding what it is like to have an experience like mortal fear or the urge to murder is almost synonymous with knowing what the experience is like, i.e., with, having had it and being able to recall it. The use of the term ‘‘understand’’ in these cases is meant to stress the requirement that the knowledge ‘goes deep,’ is thoroughly integrated into the repertoire of the respondent, and in particular that it is tied to the correct reactions. The common remark “You can’t possibly understand what it’s like to _________, if you haven’t been through it yourself” is, on the present account, an empirical claim and not a definitional one. It stresses the similarity between an experience and a perception since-e.g., there are many people whom we would never recognize (know) from a description but would have to have seen. The situation is not unlike the problem of understanding an untranslatable French phrase or idiom; ‘‘You can’t really understand that unless you speak colloquial French,’’ that is, unless you have all the syntactic and semantic connections hooked up. Interestingly, it is also like understanding a theoretical term such as the psi function. In all these cases, understanding is a cognitive state which generates certain S-R linkages, defined as producing appropriate responses to an open set of stimuli. It is natural for an ecologically oriented biologist to see this as an aspect of adaptation—we might call it cognitive adaptation. Notice that cognitive adaptability is different, being at the level of and connected with intelligence; and cognitive success (knowing an answer) is different in the other direction, being analogous to surviving for a period.

Turning to the understanding of a physical event or phenomenon, the paradigm almost all theory of explanation has been concerned with analyzing, we see an important difference between the phenomenon and the event. Understanding a phenomenon in the scientific sense requires a theory of it but the ‘‘it’’ means something which recurs in different forms and circumstances. For example, understanding the phenomenon of intermodulation distortion essentially means being able to place it in the framework of scientific knowledge which involves saying what its properties and causes are, knowing how to recognize it in all its manifestations, and this is very similar to understanding the pluperfect in Latin—the alpha, beta, and gamma criteria all apply. Understanding a particular event is a lower-level case corresponding to understanding a particular use of the pluperfect. Just as we can translate the latter description into “knowing what is meant by’’ a particular use of the pluperfect, so we can often translate ‘‘understanding a particular event’’ into ‘‘knowing what caused it’’ or “knowing its cause’’ or “knowing the reasons for it” or “knowing what kind of event it is.” But we call not as easily translate “understanding intermodulation distortion” into “knowing its cause” or anything else. The object of the understanding is much more complex, a set and not a point. It is variations in the kind of object that is said to be understood that control the interpretation of “appropriate” in the alpha criterion, and a large part of the work involved in getting computers up to the comprehension criteria has to be concerned with analysis of what understanding consists of in a given domain**.** A single context-free program for this can no more be given than it can for a computer intended to use the word ‘‘good’’ correctly, or—although this is a much simpler case—the word ‘‘large.’’

Efficiency in IPDs

We begin with a discussion of the usual basic IPDs, but artificial agents require other capacities than their storage, retrieval, and associated processing. An extra capacity must be tacked on at each end; information acquisition at the front and reaction at the back. The usual concept of artificial intelligence involves one or both of these capacities and for reasons that will not be discussed here, this is probably essential. Attempts have been made in the literature to reduce IPD efficiency measures to a single scale—time for retrieval is perhaps the best candidate—but we shall set out the desiderata somewhat more fully.

It is clear that maximizing storage capacity (1) is desirable although it is equally clear that it normally counts against cost (2), and recovery speed (3), and possibly storage speed (4), to move in this direction. Internal efficiency involves a great many factors of which the division of memory between short- and long-term (5), the division of processing capacity between central and peripheral processors (6), and the use of more than one simultaneous program (7) are three important variables to be optimized. Redundancy (8) and inconsistency (9) checks are likely to be necessary, as is fast feedback to indicate completion of storage (10). A general desideratum whose very presence indicates the necessity for compromise is reliability (11), in both the hardware and software sense, though the former will of course show up through any operational definition of the latter.

We now specify the task as any one of the set that involves handling more information, by orders of magnitude, than can be wholly stored in a usefully accessible way, where ‘usefully’ is defined by reference to the maximal reaction time consistent with the goals for the agent, e g., survival in a fight with a predator, producing a book from the stack before a freshman gives up, etc. The description ‘‘informationally rich environment” in the introduction refers to this situation. It is possible that artificial intelligence, properly so-called, *logically* requires this condition but it is certain that most applications of it will involve this condition in fact.

In such situations, desperate measures are required that will, in addition to optimizing the above eleven parameters and others, acquire the relevant material from the environment (12) and effect the most apt amongst several relevant and valuable reactions (13).

Without, for the moment, considering the last two desiderata for an interactive agent, we can see that two keys to handling (over-) rich input are selectivity in allocating storage space and use of radically economical filing systems. The main trick consists in *creating* redundancies in the input. What can be counted as a redundancy obviously depends on the utility profile for the program/agent. What makes a datum redundant may be the trivial discovery of direct or approximate duplication of the datum—watching for this is what makes the ordinary redundancy check (8) useful. But this cannot cope with a rich informational environment. To do that we have to discover extrapolable regularities in the input. If there are none, the environment is totally chaotic and cannot be handled once capacity is reached (or before). But the slightest long-term deviation from randomness creates the possibility of using the (stochastic) pattern involved to render a large or infinite amount of input redundant since it can be generated internally ruin the stored pattern-making and pattern-recognizing become the lifelines of the drowning IPD. The fundamental question for the evaluating routine is whether the mismatch between a new datum and what is already stored, or can be internally generated, is more serious than the use of resources involved in storing the exact datum. (The resources, of course, include the time arid energy costs involved in storing, as well as the space.)

The use of the term ‘patterns’ covers what we refer to in various contexts as generalizations, rules, laws, tendencies, constancies, invariants, concepts, correlations, and properties. These are the life-lines that keep us, precariously, from drowning. But possession of these is *also* the skeleton of understanding, for they provide the capacity to answer new questions, to extrapolate and interpolate, and to identify equivalences. In short, the best way to handle a rich environment is to understand it. One real-time implication of this is truistic: The only way to predict in any environment is to discover underlying regularities. What we shall here call the comprehension theorem goes beyond this, however, by linking comprehension (which is both more and less than the power to predict) with efficiency in IPDs (which involves more than the use of regularities). In particular the comprehension theorem claims that comprehension is nothing more than a dispositional capacity side-effect of an effective information storage and retrieval system in a complex, open-ended environment. An environment is open-ended if its content increases with time or closer examination; it is complex if its features require more than a *few, simple, predictable* regularities to describe it completely.

The second storage trick is modelling. Modelling is a kind of second-order or analogical patterning which uses either a physical or a formal system as time pattern, plus a conversion rule, and usually a list of qualifications. For example, the atom’s structure can be stored using the previously stored structure of the solar system as long as we remember the scale-transformation and a few non-equivalences. The simplest cases of modelling are the metaphor and the analogy itself. The must significant model is man, and anthropomorphic explanation its monument. Computers may be expected to develop a different repertoire of basic models from us although it will, by design, overlap. Thus they will correctly be said to comprehend things that we do not, or to comprehend in a radically different way some things we do understand. Correlatively, since humor is partly a useful tension-reduction device and is connected with mismatch recognition, computers are likely to have their own jokes.

The present approach might be thought instrumentalist or conventionalist. A critic might suggest that real understanding involves a *true* insight in to, for example, nature and not just a convenient device for storing data about it. We may therefore inquire into the difficulty between a mnemonic and a model. The difference is depth. A model allows, indeed encourages, interpolation and extrapolation although it may prohibit *some* directions. A mnemonic is a convenient filing device for a finite set of facts and every extension of it is illicit. The value of a model, in IPD-efficiency terms alone, is not just to facilitate recall but to make future or past observations redundant, and this it can do only to the extent it reflects reality, since reality is the source of those observations. Thus models must be realistic to yield understanding rather than the illusion of understanding—so the equivalence implied by the comprehension theorem is upheld.

A theory may be anything from a naked law or analogy to a complex of existential and universal axioms together with their development into theorems and lemmas. In either case it is covered by the previous remarks.

It must be noted that the gains in efficiency from the use of these devices are put in jeopardy by any use of extremely complicated encoding or decoding rules, lists of exceptions, or lack of reliability in the model-manipulative rules. (Moving to the model from the law already represents a loss of simplicity because the transformation rules must be stored. In practice this is more than offset by the greater fertility of time model. Also the list of limitations on a model is usually offset in practice by the need to remember the degree of inaccuracy in almost all physical 'laws.' Finally the use of models for storage does not involve learning a new pattern, which distinguishes it from many new laws.) Aging physical theories often develop these symptoms and it is hard for those whose knowledge is built around them to notice the marginal cost point where the total cost of the complications makes them less valuable to the learner than a radically different approach. The comparative youth of the great innovators in science may be more a product of their lesser experience than their greater brilliance.

We should now note that an agent which interacts with the information environment will probably go beyond the preceding devices of pattern and model. It becomes necessary to use and refine peripheral processing systems. In the IPD this distinction is not necessarily a sharp one but for various reasons of internal economy it is likely to be so arranged as to make the distinction important. On the human side it is crucial since it is at the periphery that we use the most radical information-filter of all, the concept. The origin of this in perceptual constancies, its role in the phi-phenomenon and illusions would take us too far afield here. But it is clear that just as the development of logic and rules and analogs enables the central processor to define equivalences and generate infinities of output and thus cut great swathes of redundancy through the forest of facts, so the perceptual mechanisms—both hardware and software—do the same to the environment in converting it from energy to data, or from data to information, or from information to putatively useful and usable information. Corresponding to this pre-processing there will often be a need for post-processing—evaluation and simplification, for example.

Two concluding remarks about the affective side of comprehension. It seems plausible to suppose that the necessary feedback signal announcing successful discovery of a model/law/redundancy that takes care of a new datum corresponds to the flash of insight, the feeling of understanding. And it seems plausible to think that cases of understanding a moving experience are very like modelling in the sense that matching output will be exceedingly difficult in any way that does not involve having the model already in the memory. An analogous case is the challenge to phenomenologists to show how they can dispense with the concept of a material object in favor of a phenomenological description and still produce appropriate reactions.

Applications

Outside the computer field, the most interesting applications are possibly in the philosophy of science (explanation—theory), philosophy of history (the empathy theory), clinical psychology (theories of 'clinical insight'), and philosophy of mind where the consequences for the mind-body problem and the subjective-objective distinction are significant. The details of these applications have been or are being developed elsewhere.

CONFERENCE DISCUSSION

Scriven first wanted to point out that his particular view of comprehension necessarily implicates considerations of memory as well as of inference. In addition he felt that the concept of comprehension should include not only linguistic comprehension but also the comprehension of affective and motoric behavior. A complete model of comprehension, he though, should also include provision or two kinds of "filters." One would be a "redundancy filter" that would operate on previously stored "templates" to find which one best matches a current. problem situation; to the extent that the perceiver has found a template that provides a reasonable approximation to that situation, he can be said to understand the problem well enough to function correctly in that context, and he does not need to store further data. The other kind of filter would be an "inconsistency" or error-detecting filter that would prevent the storing of contradictions.

Goodman wondered, however, how one could explain the child's acquisition of "templates" and the ability to scan them for matches with current experience. Bever pointed out that it is a mystery how children are able to arrive at linguistic generalizations (which are special classes of ‘‘templates’’), and how they are able to assign probabilities to the various possible ways of formulating generalizations from their finite set of experiences. He asked how Scriven viewed the problem of explaining the child's ability to discover which templates are relevant and fruitful, and which are irrelevant. Scriven cited the work of Minsky and Papert (1969) concerning the effect of differential reinforcement on Perceptions, as a way to approach, through computer simulation, the problem of template acquisition. But upon Bever's further probing, Scriven acknowledged that one may have to adopt the nativist position that human beings have some built-in mechanism that makes them tend to favor some dimensions over others as relevant templates against which to assess future experience.

Goodman pointed out that Scriven's comprehension theorem (that comprehension is the disposition, or capacity for effective information storage and retrieval under the constraint of the impossibility of total storage) was an interesting alternative to his own prior view that it is possible to have total a storage. Simmons commented that an implication o the notion of total storage would be that in trying to find a template to match a new situation, one would have to plow through many more of the trivial details in memory than otherwise, with the result that it might take as long to retrieve the correct template as it took to experience the event in the first place. This would make for a very inefficient comprehension system. Scriven tended to agree, suggesting that what may happen is that when we first experience a situation, we store only its main template features, in order to economize on storage space; even if we had the capacity for total storage, we wouldn't store everything because that would slow us down when it came time to retrieve the memory.

Bever raised the issue of whether a theory of tasks (which would include language performances as a special case) wouldn't prove even more difficult to formalize than the more restricted theory of language has proven to be. Scriven thought that a good reason for going the ‘‘task" route, as opposed to narrowing attention to language comprehension alone, is the potential "social payoff" of doing so. One is more likely to be successful in training people if one does it in the framework of "tasks." You should ask not whether a person "comprehends" but what a person must *do* to *demonstrate* comprehension. If you then move in the direction of training the person to perform such demonstrations, you are dealing with a manageable situation inasmuch as there is a small number of tasks that could be used its assess comprehension. In addition, one has developed a method of assessing broad comprehension that has sound training implications. This, Scriven said, was what he meant by "social payoff." He added that this is the direction that research should go; one should abandon looking into memory versus inference versus comprehension per se.

Trabasso didn't really see what this had to do with clarifying our special interest in understanding how language itself is comprehended. Scriven replied that when one narrows one's focus too radically, so that only a small sample of specially devised commands, questions, and syntactic forms are studied, this constrains the usefulness of the theories that lie behind comprehension when conceived in its broadest possible scope. Again he stated that the concept of comprehension must include not only its purely "linguistic" aspects but also the affective aspects, the comprehension of instructions, and the teaching skills. One must devise tests that inquire into the subject's ability to do appropriate things with the materials these tests contain. This may involve paying attention to the situation setting, which may be as important as the "linguistic" aspects. A satisfactory account of *real* linguistic performance will never be possible without attention to these other dimensions of comprehension which, at first sight, we *contrast* with linguistic comprehension.

Trabasso asked how Scriven would handle, within his system, the way inferences are carried out. Scriven outlined the following: the message comes in, it is 'unpacked' and something is stored depending upon its use at the moment. Behind the 'unpacking' and the storage is an analytic process which is the same device as that behind retrieval processes. Most of the future inferences that can be made depend upon how the item is stored (what templates it is stored on). As an example, suppose one is told that A is bigger than B, and B is bigger than C. This is likely to be stored as a visual model. Suppose it is. Then one is told in addition, or considers the possible conclusion, that A is bigger than C. This hits a *redundancy check* with the results that this additional piece of information is not stored—or is recognized as a valid inference—since it is already contained in the template (the visual template). Another way to state this is that the implication is stored along with the template representation. As another example, suppose you have read the five axioms of Euclid. Now should you conclude that you have stored all the theorems or not? Obviously not, but what may be stored is the general strategies that are earned concerning how to reconstruct or go about proving particular theorems. Freedle wondered how Scriven would handle the following difficulty with the two mathematical examples just mentioned. With the three-term problem, the required inference is but a *one-step* affair and the redundancy check principle is easily applied to rule out the storage of this final piece of information that A is bigger that C. However, with the Euclidean example, a problem arises when we attempt to apply the redundancy check principle. The problem is this. It is well known that the theorems of any mathematics system are tautologies consequent upon the axioms. Yes mathematicians go to a great deal of trouble to record their theorems and the proofs of the theorems even though, in a sense, the proofs are totally redundant given the set of axioms from which they stem. What does this do to the viability of the redundancy check principle? Scriven replied that while from a mathematical point of view the theorems are, truly, logically redundant, it is quite possible that we cannot claim that the theorems are psychologically redundant because redundancy because in a psychological sense is a function of what one *can* read out as an immediate consequence of what you have stored. From the point of view of retrieval, it is more practical to remember (store) the theorems simply because it is faster than having to prove them from scratch (if indeed we can prove them from scratch). The whole structure of knowledge that we are operating with, which determines what we retain and what we discard as 'redundant,' is function of how sophisticated our recovery mechanism is, and also a function of what we think the demands will be at the time retrieval is required.

Carroll asked whether the constraints due to the impossibility of total storage can be taken to imply that comprehension can never be complete. Scriven indicated that the notion of complete comprehension really is a function of contextual demand and can be likened to "complete" descriptions of a person, say. That is, if one's current purposes are motivated by just mild curiosity about someone, and one gets back a 45-minute description of who that new person is, what his interest are, etc., then one may say that this is a complete description of the new person with respect to one s current purposes. But a physician may say that this same description is quite inadequate and incomplete for his purposes. A similar point can be made about many scientific laws—they actually are only approximations to the real world (an idealization of it) which suit our current scientific needs and purposes.

REFERENCES

Minsky, M. L., & Papert, S. *Preceptrons: An instruction to computational geometry*. Cambridge, Mass.: M. I. T. Press, 1969.

Scriven, M. Dimensions of understanding. In M. Kochen, D. M. MacKay, M.E. Maron, M. Scriven, & L. Uhr, *Computers and comprehension.* Santa Monica, California: The Rand corporation, Memorandum RM-4065-PR, April 1964