Behavioral and Brain Sciences, 15, 1992, pp. 457-458.

Problem spaces, language and connectionism: Issues for cognition

Patrick Suppes
Department of Philosophy, Stanford University, Stanford, CA 94305
Electronic mail: *suppes@csli.stanford.edu*

Reading Allen Newell's *Unified theories of cognition* has been a great pleasure. It is a book one can argue with and enjoy on almost every page. It has an audacity and sweep that is rare in psychology. Much of what Newell has to say I agree with, but of course it is the function of a commentary of this sort to stress the disagreement. All the same, I shall not be able to refrain from returning at several points to things he has to say that we should all appreciate and learn from. I comment on three general issues.

1. Problem-space hypothesis. Newell emphasizes that the uniform use of problem spaces as the task representation is a central aspect of SOAR, his exemplar of a unified theory of cognition. There is much that is appealing about this, but there is also a serious problem of generality. I am reminded of the set-theoretical semantics of Montague (1974), so popular in linguistics and philosophy a decade or so ago. Without restriction, set theoretical apparatus can always be made rich enough to describe any structure likely to be encountered in language or elsewhere in nature. This is just the problem. It is as if we thought we were making progress in physics by introducing the particle-space hypothesis, that is, the formulation of all physical problems in terms of particles. We have not made significant scientific progress, however, until specific further constraining assumptions have been made and we understand how to apply them in rich detail to many different domains. I am not negative about Newell's introduction of problem spaces as a uniform method of importance (pp. 161 ff.), but I am skeptical as to whether he has at all persuaded us that this in itself has the great significance he claims. An obvious alternative, put forward only briefly in the last chapter, is that human cognition consists of many distinct modules, only a small number of which can be properly regarded as problem spaces. What Newell says about the uniform use of problem spaces is not particularly wrong in all kinds of places, but rather too many of the details are missing. I turn now to one striking example, a strong candidate for modularity.

2. Language learning. In the last chapter a fairly large section is devoted to language. There is an interesting general discus-

sion of how SOAR might apply to the learning of language, its comprehension and production. Use is made here of the important notion of chunking that Newell rightly brings to the fore throughout his discussion of many different psychological processes. On the other hand, for those who have attempted to deal with the troublesome and intricate details of language learning, comprehension, or production, what is said is at the most a very general prolegomenon to some future genuine theory of language. It is perhaps my central criticism of the book that Newell can set before us what he claims is an exemplar of a unified theory of cognition and yet provide only a very sketchy treatment of language. Many psychologists and philosophers would argue that it is the capacity for the use of language that is the most cognitively discriminating feature of human cognition as opposed to that of other animals. Newell likes to mention various well-established experimental cognitive phenomena, and this is something I commend. On the other hand, there must by now be more than 10,000 articles reporting substantial psycholinguistic data on language acquisition. There is no attempt on Newell's part to deal with any of this overwhelmingly large and complicated experimental literature. I do not mean to suggest that he should give us some kind of encyclopedic report on this literature, but I do mean to suggest that any candidate theory of cognition must come to terms with language in real detail in order to make a serious claim to being a prototheory of human cognition.

In this connection it is important to note that the list of characteristics of SOAR that "agree with the basic shape of human cognition" (p. 227) consists largely of characteristics we also expect of animals that do not command the use of language. For example, SOAR and animals that are not human behave intelligently, are goal oriented, are interrupt driven, have default behavior that is fundamentally adaptive, have recognition mechanisms that are strongly associative, do not understand how their learning procedures work, use indefinitely large bodies of knowledge, are aware of large amounts of immediate detail, and are distractible. I am not saying that these are not features of human cognition, but they are also features of animals that have no use of language. I think it is too bad that Newell did not make a strong effort to differentiate cognition that depends on language from that which does not. It is ironic that the most obvious and distinctive symbol-processing mechanisms of human cognition, namely, those of language comprehension and production, receive so little attention from the premier proponent of the symbol-processing approach to cognition.

3. Connectionism. Newell acknowledges in his introduction (p. 39) that the rising tide of connectionism may sweep over all of cognitive science. I am skeptical of this, at least of connectionism in its present form, as I am equally skeptical of any purely symbolic processing approaches. What is really too bad, however, given Newell's unique position in developing the importance of symbol processing in cognition, is that he did not meet head-on the current onslaught of connectionism with a full-scale intellectual analysis of how he sees the relations as they now stand between connectionism and symbol-processing systems and where the theories of cognition of the future will go. An aspect of this that is important is that connectionism itself has clear roots in the stimulus-response theory of earlier behaviorism, the kind of theory that Newell rejects and of which he is in various places rightly critical, but the criticism of the excessive simplicity of past stimulus-response theory can also be leveled at purely symbolic processing theories of cognition of the kind he advocates. Newell gives on pp. 484–88 a tantalizing comparison of SOAR with current work in connectionism; hence it may be considered a rather lame criticism on my part to ask for something the author did not offer, a much more detailed comparative analysis. My excuse is that he is much better prepared than most of us to give such an analysis.

This brings me to a final point that is an outgrowth of reflection on the relation between connectionism and SOAR. Central to connectionism is the fundamental role of probabilistic considerations in all animal and human learning and performance. Except for the simplest sorts of highly structured tasks, probabilistic considerations enter at every level, in uncertainty faced in perception, in uncertainty faced in planning, in uncertainty faced in task execution, and in uncertainty faced in learning and change. In working in the framework of computer programs, as is the case with SOAR, it is very tempting to leave too little room for these pervasive probabilistic features of both the world we live in and our methods for dealing with them. Certainly Newell mentions such probabilistic aspects of SOAR's performance here and there but SOAR does not have a central place for the probabilistic features that are so pervasive in the world and in actual behavior. In particular, the fairly lengthy discussion of preferences and decision procedures (pp. 170–74 and elsewhere) takes surprisingly little notice of the large literature on probabilistic theories of decision making in psychology, which are so naturally congenial to connectionists. The brief mention on pp. 435–36 is written as if pertinence to psychologists of the concepts of subjective probability and uncertainty issued mainly from experiments of relatively limited interest on "choices among abstract gambles" (p. 436). To paraphrase Svengali (as played by John Barrymore), in the bright lexicon of Newell's theory of cognition there is no such word as Bayes.