Cohen, on the other hand, was concerned with explaining subjects' apparent failures to use base rates properly (as in the famous "cab color" inference problem, pp. 328–29).

Of Anderson's four cognitive domains, three involve commonplace, learnable, and purposeful behaviors: classification, causal inference, and planning. If the environment affords any strategies for these tasks that frequently produce respectable results at low computational cost, then research of the sort described by Anderson would seem to have a high a priori prospect of success.

People will presumably do something when asked by experimenters to perform in these task domains. It is plausible that what the subjects might choose to do would involve some trade-off between effort and efficacy.

A search for a descriptive model of those subjects' trade-offs, like the one outlined in Table 1 of the target article, would thus seem to have bright prospects. There are plenty of degrees of freedom available to model makers. What shall we designate as the measure of success or of cost? Which environmental constraints shall we recognize in the model? Which simplifying assumptions shall we adopt?

Success in coming up with a model seems unsurprising. Lack of potential surprise, rather than criticism about iterative hypothesis refinement that Anderson easily deflects, dilutes the evidence's probative power; for as probabilists (Polya 1954) and nonprobabilists (Shackle 1949) alike remind us, we learn little from unsurprising events. That is, our prior beliefs about the rationality of human cognition are apt to be unchanged by modeling success that was never much in doubt in the first place.

The properties of memory are presumably not as open to subjects' choices as the other behaviors. Yet, here too, the modeler's jacket is generously cut. We find in the conclusions of the target article yet another degree of freedom: Which "components" of the "cognitive system" are to be modeled as locally optimal? (By the way, in the example task of phone number recall discussed in the conclusion, might not the "cognitive system" learn a few mnemonic tricks if remembering phone numbers were important enough?)

And yet, of 21 phenomena studied by Anderson, only seven are accompanied by a remark that suggests that other workers interpret the phenomenon as evidence of irrationality. In the other 14, the rational analysis illuminates issues other than the "big question": Is cognition adaptive?

In the three phenomena cited in connection with the "history factor" of memory, a beautiful, nonobvious similarity is established among diverse engineering problems of designing information-retrieval systems. The six categorization phenomena support a critique of existing research in the field: a misplaced (in Anderson's view) preoccupation with predicting category labels. Phenomenon 16 links results from the causal inference and the categorization models to suggest new hypotheses for investigation. The final four phenomena support principled criticism of the kinds of tasks often studied in the planning and problem-solving literature.

In addition, for all four domains, Anderson motivates interesting Bayesian models of how to approach a task. Quite apart from whether or not people are "rational," these models are surely interesting in themselves, and useful to the artificial intelligence enterprise, among others.

Anderson's rational perspective, then, is a successful point of departure for science. Existing interpretations of experiments are challenged, new hypotheses are suggested for further experimentation, and similarities are disclosed among problems that otherwise appear distinct. A powerful general-purpose modeling tool, Bayesian analysis, is artfully adapted to specific applications.

Does the worth of Anderson's perspective depend much on whether or not human cognition is really rational? It had better not, since on the evidence of the Cohen & Kyburg commentary, the experimental record is equivocal. People don't even always agree on what being rational means.

As a springboard for successful science, the opposing perspective is fruitful, too. Systematic inquiry into how behavior departs from selected normative accounts of rational behavior has been a fertile source of inspiration for hypotheses (in the work of Kahneman & Tversky, 1979, for instance) and cognitive engineering design (such as the extensive work in the nonprobabilistic management of uncertainty surveyed by Prade 1985).

Big questions fascinate us; for many scientists, interest in such matters is surely part of the explanation of how they came to be scientists in the first place. Having a suspicion about how the answer will turn out can be a strong motivating force. We all know good scientists whose work is informed by deeply held religious convictions, and other good scientists whose belief in a universe without any hint of the supernatural is just as strong.

If science is roomy enough for both views on that particular big question, then surely rationalists and their critics can both be accommodated. Professor Anderson succeeds in demonstrating that his beliefs are respectable. On the other hand, it is probably too much to hope that science will resolve the rationality question, any more than it will settle questions of religion. There is other work to be done, on questions that are smaller, but whose answers are more nearly within our cognitive grasp.

Computational resources *do* constrain behavior

John K. Tsotsos[1]
Department of Computer Science, University of Toronto, Toronto, Ontario M5S 1A4, Canada
Electronic mail: *tsotsos@ai.toronto.edu*

Anderson says: "A rational approach encourages us to inquire about the structure of our actual environment and to design an algorithm optimal for it rather than to design algorithms which would only be optimal in some bizarre world." The rational world Anderson proposes is as bizarre as the approaches he is criticizing. Behavior is not only a function of environment; behavior develops as a satisficing function constrained by environmental conditions as well as computational resources (and perhaps other things). In fact, given our current understanding, the limits on computation imposed by our brains may play the largest role in shaping behavior.

Anderson says that in his work he has yet to find computational limitations posing danger to his scheme. Unfortunately, I don't think he has looked hard enough. The combinatorial problems are very apparent, and in fact in most (if not all) natural problems, optimal solutions are computationally intractable in any implementation, machine or neural. A few examples are in order.

(1) Vision
- Unbounded visual search, using a passive sensor system is NP-complete (Tsotsos 1989; 1990a)
- Unbounded visual search, using an active sensor system is NP-complete (Tsotsos, submitted)
- Polyhedral line-labelling is NP-complete (Kirousis & Papadimitriou 1985)

(2) Reasoning
- Finding the optimal satisficing strategy for simple and-or graphs is NP-hard (This refers to the task of deciding which operator to use to reduce a goal to its subgoals.) (Greiner 1990).
- Finding the best explanation for a class of independent problems using probability theory (and several other

forms of abduction) is NP-hard (Bylander et al. 1989).
Abductive reasoning for all but the simplest theories is NP-complete (Selman & Levesque 1989).
Many forms of default reasoning are NP-hard (Kautz & Selman, in press; Selman & Kautz 1990).
Many of the strategies for defeasible inheritance in taxonomic hierarchies are intractable (Selman & Levesque 1989).

(3) Neural networks
For directed Hopfield nets, determining whether a stable configuration can be found is NP-complete (Godbeer 1987).

This listing only scratches the surface of the literature on the topic; there are many more examples. As should be clear, the problem-solving type of task, where a sequence of actions is required as opposed to single actions, is not the only one that has a combinatorial nature, as Anderson claims. Neither are the problems above obscure and isolated; rather, they are quite broad and natural. All "interesting" intelligent problems appear to be susceptible to combinatorial explosion. It is important to stress, however, that the examples given above do not by themselves "prove" that these problems or cognition in general are computationally intractable. They simply constitute evidence that the computational issues are real and may place severe constraints on algorithms proposed for the problems of cognition.

What does a computer scientist do when confronted with such a potentially intractable problem? A variety of approaches are possible.

(1) Develop an algorithm that is fast enough for small problems, but would take too long with larger problems. This approach is often used when the anticipated problems are small.

(2) Develop a fast algorithm that solves a special case of the problem, but does not solve the general problem. This approach is often used when the special case is of practical importance.

(3) Develop an algorithm that quickly solves a large proportion of the cases that come up in practice, but in the worst case may run for a long time. This approach is often used when the problems occurring in practice tend to have special features that can be exploited to speed up the computation.

(4) For an optimization problem, develop an algorithm which always runs quickly but produces an answer that is not necessarily optimal. Sometimes a worst case bound can be obtained on how much the answer produced may differ from the optimum, so that a reasonably close answer is assured. This is an area of active research, with suboptimal algorithms for a variety of important problems being developed and analyzed.

(5) Use natural parameters to guide the search for approximate algorithms. There are a number of ways a problem can be exponential. Consider the natural parameters of a problem rather than a constructed problem length and first attempt to reduce the exponential effect of the largest valued parameters.

NP-completeness effectively eliminates the possibility of developing a totally satisfactory algorithm. Once a problem is seen to be NP-complete, it is appropriate to direct efforts towards a more achievable goal. In most cases, a direct understanding of the size of the problems of interest and the size of the processing machinery is of tremendous help in determining which are the appropriate approximations. Could evolution have discovered this through millennia of experimentation?

It would be an extreme and untenable position to claim that behavior is made up of a large number of side-effects due to approximations; however, how much of our behavior can be legitimately put into this class? This is currently unknown and seems to me an interesting question for further study.

NOTE

1. The author is also affiliated with the Canadian Institute for Advanced Research.

Human and nonhuman systems are adaptive in a different sense

Tamás Zétényi
Department of General Psychology, L. Eötvös University, H-1378 Budapest, Pf.4, Hungary

In recent treatments of conceptual phenomena by cognitive science, three levels of analysis have been distinguished: a computational, algorithmic and an implementational level (see Marr 1982). In his target article Anderson undertakes to perform all three.

Anderson argues that if we describe the statistical structure of the environment, we can predict human cognition, which is supposed to be an optimal response to it, and is therefore adaptive.

He attempts to support the above claim by analyzing the behavior of nonhuman information-retrieval systems such as libraries or computer data bases. The results of the analysis of these systems are indeed correctly predicted by the carefully selected list of phenomena. What we need is an algorithm to represent the possible transformations between input and output. Bayesian theory is undoubtedly suitable for this purpose, so if we try to implement the resulting algorithm into another physical system, namely a connectionist(like) network, it is likely to run without a single error message.

What I find problematic in this is the claim that the memory functions of a nonhuman information-retrieval system can be plausibly extended to human memory. In spite of the cases provided by Anderson I wonder whether the algorithm can be implemented into human cognition.

Libraries and psychological experiments are indeed similar to each other in some respect. They have a limited set of items to handle. There are similar tasks too: A book or an item of a nonsense trigram list must be remembered upon request. It is true in both cases that memory performance will increase with practice and decrease with duration; and frequency also has an effect on it.

But let us assume that items in a library or files in a data-base are not independent entities; they communicate just as they do in cartoons. They can retrieve certain ideas, form categories, make decisions, and solve problems without any assistance. If this were the case, their behavior would indeed be similar to that of humans. This is the type of performance we call cognitive.

Nonhuman information-retrieval systems, as far as I know, never make changes in the entire material without human assistance. Subjects in a psychology experiment try to answer according to the instructions provided by the experimenter. Nevertheless, they show a strong tendency to retrieve items which are not included in the list. The types of errors are covered in detail by textbooks on human memory. Retrieving out-of-list items is excluded in the case of nonhuman systems. For example, we will not find on a library shelf a monk copying some codex, although these two are commonly associated with the concept "book" in memory. Similarly, it does not happen that a volume not satisfied with its borrowing index changes the color of its cover. It does not occur, either, that two different files of a data-base exchange strings without assistance. Books have never moved to certain places because they decided to form a new club (categorization), or because they found the shelf wet (reasoning). These would be instances of adaptation; these answers are certainly optimal with respect to environmental circumstances by some standard. The basic structure of a library or data-base will not change over time, so there is room for rational analysis. In contrast, a human retrieval system is an ever changing "environment."

Note that there is something strange in the argumentation above, namely, personification. I have personified the book or