HIGH-PERFORMANCE COMPUTING AND ITS PROMISES FOR THE BEHAVIORAL SCIENCES

Chaired by Lynne K. Edwards, *University of Minnesota*

Perception, models, and data: Some comments

PATRICK SUPPES
Stanford University, Stanford, California

Some comments are offered on the papers given at the conference, which are divided into three groups: visual perception, models and neural networks, and data analysis. The analysis stresses the complexity faced by scientific theories in each of these three areas, and consequently why the demand for computing capacity will continue to increase, with no practical bound in sight.

I have divided the papers on which I comment into three groups: visual perception, models and neural networks, and data analysis.

VISUAL PERCEPTION

I put the four papers on visual perception in the order in which we heard from James Cutting (1997), Daniel Kersten (1997), Mary Kaiser (Kaiser & Montegut, 1997), and Jeffrey Mulligan (1997). These are the most hopeful four papers I have heard in psychology in a long time. They show how many things there are under way and yet to be done.

Anybody can see a need for high-performance computing in psychology by considering how much of the cerebral cortex is taken up by the visual system. It is a complicated, big system, from a computational standpoint. There will be no end to what we want to try to do. One concept I want to introduce in reflecting on James Cutting's paper, but on the others as well, is this: We should think of the processing in the visual cortex as being computationally irreducible. It is a nice concept that was, in fact, introduced by the father of Mathematica, Steve Wolfram (1985). He introduced it in the context of cellular automata.

If we take, for example, the beautiful classic case of two bodies interacting by gravity, then we have a wonderful computationally reducible theory. We do not have to wait to see where the bodies are going to go (e.g., the sun and the earth). We have a good theory that predicts accurately, and we can compute very quickly. This case is computationally reducible. But, for many cellular auto-

mata, we are not going to do better than run the system. The kind you actually run is a direct simulation of cellular automata, in order to know what it is going to do, what the configuration looks like after, say, 10,000 trials. The conjecture is that this may be true of the visual system in many essential ways. I will discuss this point more later in this paper.

This conjecture implies that we will have an insatiable appetite for computing. There are some differences here about how much we believe in extensive computing. It seems to me that these four papers, each in its own way, demonstrate very well the extraordinary desirability of having high-performance computing in psychology, if you are concerned with the details of visual systems at all, in any kind of elaborate way.

Let us compare the computational demands of mathematical proofs. From a computational standpoint, the mathematical cases are all rather trivial. For example, we all know something about the proof of the four-color theorem, which involved about 1,500 cases. In spite of the extremely large number of cases that had to be analyzed, it is simple stuff relative to what we heard about here. And I think that is characteristic of the trend of much mathematics. If you think in terms of deductive structures, what we can put into solid proof form is relatively simple. The thinking involved in even the most elaborate mathematical proofs, from a computational standpoint of reproducing them, at least, is quite simple relative to the performance of the visual system.

Another related remark about this insatiable appetite for computing is this: Because we all know that there are very few agencies that will buy supercomputers for psychologists, the best hope is that consumer demand and business demand will continue dynamically to drive the market, so that every psychologist concentrating on the visual system will have good computing equipment in

Correspondence should be addressed to P. Suppes, Department of Psychology, Ventura Hall, Stanford University, Stanford, CA 94305-4115 (e-mail: suppes@ockham.stanford.edu).

the not-too-distant future, as a by-product. For example, we could have "Mars" as a video game, which would be a way to increase funding for NASA.

I was struck by Jeff Mulligan's (1997) talk about eye movements, which I have had some research interest in. The enormously elaborate computations you need to go through in analyzing eye movements remind me of what people go through in the numerical study of turbulence. It is often thought that we do not have anything in psychology that is comparable in its demands for computational power to what is needed in physics. I do not think that this is at all true. In fact, vision alone provides all the examples we shall ever need.

One of the reasons that I think there will be such a demand for extraordinary computational power is that it is unlikely that we will have an economical mathematical theory of the visual system. We will not get a theory—that can be written down in two pages—of the basic assumptions that cover most of the phenomena, once the boundary conditions are expressed. We will not have a finite set of differential equations that we can list on a couple of pages that will do the job. I think that, in many ways, Cutting (1997) and Kersten (1997) illustrated very nicely why this economical theory is very unlikely. Take Cutting's example of all the different ways in which we use cues to compute information about depth in paintings. It is a real hodgepodge of stuff, even though much simpler than the real world. Something as simplified as a painting is still enormously complicated, and not easily thought about in elegant mathematical theoretical terms. Too many different kinds of things are going on. I would suggest—even though I can be wrong about it—that we do not want to have the mistaken hope about the visual system, or about the brain in general, that we are going to end up with a unified, simple mathematical theory. There is good reason to think that is not going to be the case. We have to think about it theoretically in a different way. So, I found these papers very suggestive. They gave me much to think about.

MODELS AND NEURAL NETWORKS

The second group also gave me much to think about. Let me begin with some remarks about the papers of Richard Shiffrin (Shiffrin & Nobel, 1997) and Richard Golden (1997). I cannot resist, given what we are tackling in the visual system, posing a question to Dr. Shiffrin. Suppose we take the same attitude toward memory as toward vision. This means that we want to be able to deal with many natural phenomena. The question is, how much of memory do you think we can simulate, and is it a worthwhile scientific idea to have something to think of in terms of memory comparable to the way we think of vision? So, if we take the kind of examples given here—whether it is Kaiser's (Kaiser & Montegut, 1997) example of Mars, Kersten's (1997) complicated shadow phenomena, or Cutting's (1997) examples of the ways in

which painters produce the sense of depth—can we equal them in memory phenomena?

I think that what Richard Golden had to say about text analysis is the right level for much of the work we can hope to do in the reasonable future. Those of you who have not looked at problems of text analysis might say, "Why worry about, say, the episodic and causal structure? Let's just go for a really rich complex model of text comprehension, so that, whatever it is, pass it to me and I will comprehend it." This turns out to be a real problem. Maybe it is not quite as bad as the vision problem, but it is certainly very bad. So, we need to make piecemeal starts at analyzing text.

Now, what is interesting in areas close to psychology—in linguistics, for example—is that current text analysis work is very promising. In certain respects, Chomsky was a bad influence in being very skeptical and extremely critical of what had been the prior statistical tradition in linguistics. Linguistics had been very much behaviorally influenced by psychology and very much behaviorally oriented in the period prior to Chomsky's work in the late 1950s. A good tradition of statistical analysis of linguistic data, after a hiatus of 20 years or so, is back in full force. Those of you who follow such things know that there are massive statistical efforts now in a linguistics framework, with new attitudes and new ideas about how to deal with the complexity of texts. They are very much in the spirit of things that we are hearing about here. One big example is machine translation. A nice example comes from Canada, where legal proceedings in both English and French are regularly published. So Canadians have thousands and thousands of matching pairs (English–French) of complicated sentences and phrases to look at. There are many other examples in which the focus of attention is very large scale corpora. And, in general, beyond machine translation, there is a desire to be able to parse and understand large texts. Of course, these efforts are far from being fully successful. But, on the other hand, it is quite clear that many new things are happening and much progress has been made. It is also quite clear that, as work continues, just as in the case of the vision research, an insatiable appetite for computing power will continue indefinitely.

I have recently been involved in machine learning of natural language, and I think the biggest criticism is that we have not shown how to scale up to larger problems. The work has been in a robotics context or that of physics word problems. The criticism is that the training sets of sentences are too simple. They represent too small a piece of a language. It is quite clear that when we scale up, our attack on machine learning, with the fundamental assumptions about learning that we use, will get into massive computations very quickly.

One of the surprising things that I cannot help mentioning here is that, where computers continue to play a big role in various forms of machine learning, most of the players now are not behavioral scientists. We have an

interdisciplinary seminar at Stanford in machine learning that meets every week. It is interesting to see what the composition of that seminar is. When I started working in learning, back in the 1950s, a learning seminar had only psychologists and maybe one or two outliers from somewhere else. In this seminar, I am a sort of renegade psychologist, and I am really the only one present who has any history of working in psychological theories of learning. We have computer scientists and mathematical statisticians. We also have people from the medical school interested in DNA sequencing. I find it wild and wonderful that learning is no longer a disciplinary subject that simply sits in psychology.

Now, I have a question for Sam Williamson (Gençer & Williamson, 1997). I have also heard, as a part of this symposium and some other seminars at Stanford, about brain scans, PET, MRI, older EEG work, and now MEG. My question is this: When will we be able to get a glimpse of a real computation going on? It is clear that blood flow is not going to give us a very good insight into the details of the computation. I am not quite able to detect, from Williamson's presentation, what his secret ambitions are for the magnetic fields. I suppose my question is, "Do you think that if you work hard or your successors work hard, you will be able to really say something subtle about the actual computation? Can the activities that you can observe in the magnetic fields give us a serious representation of the computation?"

Jim Anderson (Anderson & Sutton, 1997) had endlessly provocative things to say. I want to pick up on one of them. Half jokingly, I assert that Chomsky has it wrong about language. It is not language that is innate. It is arithmetic. In fact, arithmetic is really easy compared with language. It is a simple computation that elementary-school children around the world get at most 20,000 exercises in 5 or 6 years of school. In contrast, by the time a child enters the first grade, he or she has been exposed to something like 2 million utterances—two orders of magnitude more than the exposure to arithmetic. So, it must be arithmetic that is innate in order to be learned with such relatively little exposure. Of course, I am not serious about this argument taken straight. Locke liked to think of the mind of the child as a tabula rasa. In this respect, Chomsky is much closer to being right than Locke.

DATA ANALYSIS

I classify the last two talks as data analysis in a broad sense. Albert F. Anderson (1997) gave us a detailed picture of something very important: how to really get something done in a reasonable amount of time. And Jim Ramsay (Ramsay, Heckman, & Silverman, 1997) had a lot of useful ideas about ways to facilitate what he calls "smart smoothers." These constitute a very important topic.

I also found both of these talks hopeful. If you read the wrong journals, you can think that the situation for analysis of complicated phenomena is reasonably hope-less. I will give you an example. I have in mind the *Annals of Probability*. Most of you probably do not often look at this journal, and neither do I. But I do look at it periodically. If you look at it, you may well conclude that a great deal of the world's phenomena need to be thought of as stochastic processes. Human visual processing is an example. Such processes are continuous in space and time. So, they are very complicated processes and are very difficult to analyze.

Williams (1979) has written a beautiful book on diffusion processes and Markov processes. He remarks that one of the sad things about the subject is that it has gotten so extraordinarily technical in the past 25 years. If you look at many of the papers that are written in this framework or at Williams's book, you may ask, "What could I actually use that this research has provided that is interesting from a scientific standpoint?" In many cases, it is a fairly small yield (e.g., various features of the motions that are unlikely to be of great interest). So, Albert Anderson's and Jim Ramsay's talks are wonderful therapy for thinking that the world is just too complicated. Albert and Jim are good at getting around such pessimism.

Physicists are also good at getting around this kind of pessimism. They have the attitude that things have to be done, even if the standard of rigor would not satisfy the *Annals of Probability*. To do that, we might need another 100 years to solve serious problems in quantum field theory. What is extremely important in thinking about the application of computational methods is to recognize that it is essential to roll up our sleeves and approximate in various ways—ways that can still be highly informative about processes that may in detail be unbelievably complex.

It may sound as if I am talking out of both sides of my mouth. Earlier, I was saying that I do not think that we will reduce the visual system to an elegant, relatively simple mathematical theory. Well, these two things go together. I have a sermon that I will give in about 30 sec about the misunderstanding of physics sometimes found in the behavioral sciences. You may think physics is a beautifully clear subject founded on firm concepts. It is not true at all. Physics is a wonderful chaos of thought. As the Russian mathematician Manin (1979/1981) says, "Physics will provide everything you want except clarity." If you want clarity or rigor in a mathematical sense, you will not find it. What physicists are wonderful at is picking the pieces to put together a "picture" to fit a particular problem. They do not say, "Well, I have these axioms, and I will prove this theorem about this micro behavior, for example, this superconductor property." But they understand very well how to use their marvelous ability to pick just the right package of facts and assumptions for a given problem, using just the principles needed for the given problem. Also, when an experiment is done, the analysis has very much the same flavor. So, to anybody who has the mistaken idea that physics can be reduced to a simple set of principles, I say, "Look at the publications of the American Physical Society in any

current area of physics and try to reduce it to something that has any kind of simplicity to it." So, my thought is that Albert Anderson and Jim Ramsay are proposing tools that we can really use. They provide an approach that the physicists are very good at. We need more of that in the behavioral sciences.

I just want to end now with one final remark. Albert Anderson began with a remark about our table conversation. Let me also say something. I had a clearly mistaken idea—a guess, I should say—about the population of Minnesota. I thought it was less. I was working from the population of Minneapolis, so it seemed to me it should be smaller than it is. Albert pulled out his handy-dandy portable and had the data there to look it up. We are able to settle the population question very quickly. In a nice leisurely dinner, it is easy to cover a lot of ground in a conversation. But, unfortunately, most of us do not remember details of what we previously learned. So, the dinner talk is often rather vague. I am not quite sure about this, but what we can hope for in the computing of the future will be more like what Albert did. We will say, "Hey, we've got a fast hookup here, let's just test those ideas quickly." We will get a quick response. Not meant to be definitive, not publishable, but look what we can do in a minute and a half. And so the nature of academic table talk, if we all come equipped with our connections or hand-held devices, will change the nature of the talk. The slogan is this: Speed will change the world. And I think seriously that will be the case. I want to remind you of that model often stated of learning and wisdom, ancient Athens. Actually, very few people could read—a small percentage, probably less than 10% of the citizens. At our dinner table of five, we had 20% with computing power (i.e., Albert). I think the future will be dinner with computational power and good wine

widely distributed. Dialogue and disagreement will continue to flourish, but in an ever more fertile field of fact and data.

REFERENCES

ANDERSON, A. F. (1997). Application of high-performance computing to the management of social science and demographic data. *Behavior Research Methods, Instruments, & Computers*, **29**, 86-98.

ANDERSON, J. A., & SUTTON, J. P. (1997). If we compute faster, do we understand better? *Behavior Research Methods, Instruments, & Computers*, **29**, 67-77.

CUTTING, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, & Computers*, **29**, 27-36.

GENÇER, N. G., & WILLIAMSON, S. J. (1997). Magnetic source images of human brain functions. *Behavior Research Methods, Instruments, & Computers*, **29**, 78-83.

GOLDEN, R. M. (1997). Causal network analysis validation using synthetic recall protocols. *Behavior Research Methods, Instruments, & Computers*, **29**, 15-24.

KAISER, M. K., & MONTEGUT, M. J. (1997). Of red planets and indigo computers: Mars database visualization as an example of platform downsizing. *Behavior Research Methods, Instruments, & Computers*, **29**, 48-53.

KERSTEN, D. (1997). Inverse 3-D graphics: A metaphor for visual perception. *Behavior Research Methods, Instruments, & Computers*, **29**, 37-46.

MANIN, Y. I. (1981). *Mathematics and physics* (A. Koblitz & N. Koblitz, Trans.). Boston: Birkhauser. (Original work published 1979)

MULLIGAN, J. B. (1997). Image processing for improved eye-tracking accuracy. *Behavior Research Methods, Instruments, & Computers*, **29**, 54-65.

RAMSAY, J. O., HECKMAN, N., & SILVERMAN, B. W. (1997). Spline smoothing with model-based penalties. *Behavior Research Methods, Instruments, & Computers*, **29**, 99-106.

SHIFFRIN, R. M., & NOBEL, P. A. (1997). The art of model development and testing. *Behavior Research Methods, Instruments, & Computers*, **29**, 6-14.

WILLIAMS, D. (1979). *Diffusions, Markov processes, & martingales: Vol. 1. Foundations*. Chichester, U.K.: Wiley.

WOLFRAM, S. (1985). Undecidability and intractability in theoretical physics. *Physical Review Letters*, **54**, 735-738.