



AAAI-98 Presidential Address

The Importance of Importance

David Waltz

■ Human intelligence is shaped by what is most important to us—the things that cause ecstasy, despair, pleasure, pain, and other intense emotions. The ability to separate the important from the unimportant underlies such faculties as attention, focusing, situation and outcome assessment, priority setting, judgment, taste, goal selection, credit assignment, the selection of relevant memories and precedents, and learning from experience.

AI has for the most part focused on logic and reasoning in artificial situations where only relevant variables and operators are specified and has paid insufficient attention to processes of reducing the richness and disorganization of the real world to a form where logical reasoning can be applied. This article discusses the role of importance judgment in intelligence; provides some examples of research that make use of importance judgments; and offers suggestions for new mechanisms, architectures, applications, and research directions for AI.

It's a great pleasure to address you all. I have viewed this as a unique opportunity to examine what I think about AI as a field and to say what I think to you without having to run it by reviewers. I'd like to talk about what I think is really important in the field and in particular what we're doing that's important and what's important that we aren't doing, with more emphasis on the latter. Roughly what I'll cover today is the following: first, the good news, briefly. AI really is doing well, and I'll try to give a brief overview as I see it. Then I'll digress to ask some basic questions about intelligence: What are brains really for? How did they evolve? Why do we have them? I'll argue that brains and intelligence are rooted in matters of life and death, important issues, things that people regard as juicy, things that people really care about. And then I'll move to the bad news: Assuming that its goal is to really

understand and model intelligence, AI is neglecting many essential topics. I'll talk about specific issues in perception, deep understanding of language, creativity, and logic. I'll make some suggestions about what we as a field should do based on what I've said. If I don't offend everybody at least a little, I've failed.

First the good news: AI is thriving. Unlike earlier times, AI encompasses a broad range of areas and topics. If you look through the current AAAI conference proceedings, I think you'll see this. The narrow focus on logic alone is gone. When AI people make predictions, which of course we do and should do, they're much more muted; I don't think we're going to get in trouble for the kinds of predictions we're making today. There's a significant industrial effort and many applications. Membership in the organization is stable now. It had dropped for a long time—from 15,000 during the expert systems craze—but it's now been around 5,500 for several years and maybe has even gone up a little bit. If you count institutional memberships, it's more like 7000. We have very high standards for paper acceptance. In this conference, the acceptance rate was 32 percent, and journals in AI are strong. Advances in computing hardware have been a huge help to AI. For instance, in the practical speech recognition systems of today, there aren't many new ideas compared to what we knew in the seventies. (There are a few: For example, hidden Markov models (HMMs) are important and new since the seventies.) But overall, much of the credit for making AI become real in this and other cases should go to computing power increases. The earlier ideas were adequate, but the computing power was inadequate.

There's a very impressive range of very interesting AI applications. Rather than running through a long list, I'll point you to the IAAI

...
*understanding
 cognition—
 understanding
 the way the
 brain works,
 what the
 mind does—
 is a key
 scientific
 grand
 challenge.*

proceedings this year and other years. Many grand challenges, first stated many years ago, have been achieved. We have truly autonomous vehicles such as Dean Pomerleau's NO HANDS ACROSS AMERICA project, the autonomous vehicle that drove from Washington to San Diego, with all but 30 miles hands off, and SOJOURNER, which wasn't very smart but certainly made a big splash and will be followed by more interesting systems. In automatic gene finding, AI techniques were used by Steve Salzberg to find the Lyme disease and syphilis genes recently. There have also been the automatic classification of galaxies and stars and winning game programs.

In addition, AI has a very impressive record of spin-off ideas: Lisp, SCHEME, PROLOG, maybe SMALLTALK, garbage collection, object-oriented programming. Lots of programming ideas came out of AI labs; so did the mouse, the graphic user interface, software agents, data mining, telerobots, metasearching, PCs, laser printing, time sharing, symbolic math aids, computer games, graphics, and others.

Many important future applications will require progress in AI: high-quality translation; *high-precision web search*, really finding just what you want as opposed to finding a list that hopefully contains what you want; *content-based image matching*, finding the image you want; intelligent robots, intelligent environments, smart offices, and houses; and autonomous space systems.

Moreover, I think it's very clear that understanding cognition—understanding the way the brain works, what the mind does—is a key scientific grand challenge. I think that within the next 20 years, we will have available affordable computing power sufficient to match human abilities, at least in principle. We'll have a lot more knowledge of the brain and mind. So the future looks very rosy in that regard.

Here is a chart, with thanks to Hans Moravec (figure 1). You can see that computers are very quickly approaching the chimp-human-whale cluster on the upper right. And it won't be too long after that before these creatures will be overtaken if trends continue. And there's very good reason to believe that trends will continue.

What Are Brains For?

That was the good news. Let's step back and look at the Big Picture: What is the brain really doing, what is intelligence about? This is a personal view, but I feel pretty confident and

strong about it. The central phenomena of intelligence are tightly related to life and death issues. We have brains because they really made a difference to human survivability as individuals and as a species. What exactly is AI trying to study, understand, and explain? The study of intelligence is an ill-posed problem: What is intelligence? A better question is *What is intelligence for?* That one I think we have some chance of answering. And the answer to it is that any kind of intelligent phenomenon that we see, whether physical or behavioral, is there because it really serves the organism's survivability purposes. Brains are for matters of life and death. Brains are primarily organized to perform fast action selection in order to promote survival. Brains make decisions on issues such as fight or flight, on the finding of food and water, on the selecting of mates and reproducing, and in only a slight exaggeration, I think the basic operations the brain is doing is asking over and over again, quickly, *Can I eat this, can it eat me?* Now, it doesn't seem to us that this is happening. Partially this is because given the degree of intelligence and development that we have as a culture, we've created a world where food is plentiful, and predators or dangers are kept at bay. But that doesn't tell us much about where the brain came from. We surely haven't changed very much physically in 5000 years. Five thousand years ago the world was a much more dangerous and uncertain place. We've been able to relax about these life and death issues because we've created a supportive, protective culture. But I think that the brain arose during much less secure times.

How did the brain actually come to accomplish fast action selection? It accomplished it through the evolution of a lot of critical structures: innate drive, reward, and evaluation systems. We know what we like. We know when something we experience is good or bad. We don't have to be taught that; in fact, that's the mechanism of teaching. The brain consists largely of many parallel, short, interacting chains of neurons that connect attention and perception systems to action systems. We know the chains are short because we can act on things within roughly 100 milliseconds, and given the latencies of neurons, which are on the order of a millisecond, there can't be many more than 100 layers of neurons between inputs and outputs. That's very short! There's not much chance for backtracking or serial reasoning. Of course, the paths could be very broad and bushy: There could be many parallel paths and many neurons involved, but the depth can't be very great.

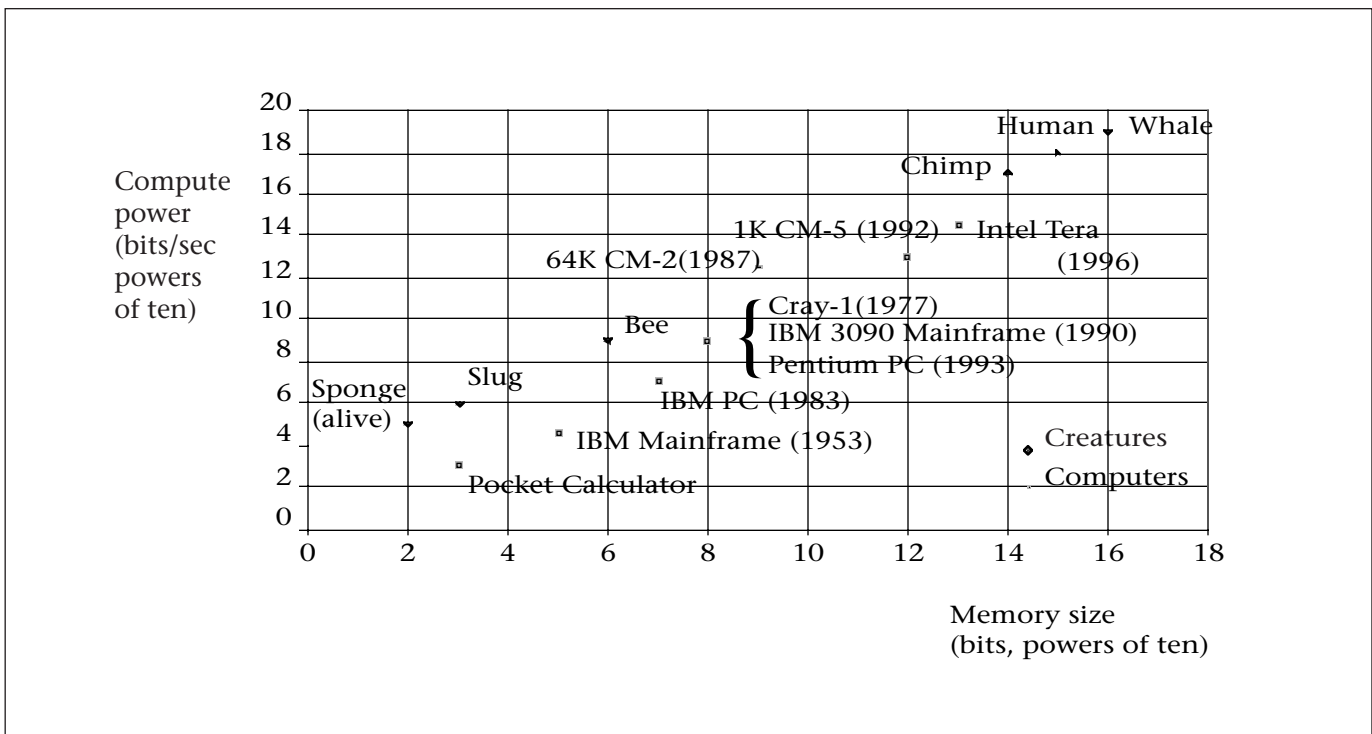


Figure 1. Relative Compute Power and Memory.

The other very important things we have are many metasystems that watch other systems, that in particular watch the actions of these short chains (and other parts of our brains) and look for regularities that they can predict. They can generate shortcuts—essentially predictive mechanisms—and also try to generate metaobservations and learning controls. This is very important for intelligence and probably one area in which people are clearly differentiated from other creatures. There’s been a lot of discussion over what drives learning. One popular theory is that learning is failure driven. I think that’s roughly true, but I’d like to broaden it a bit: I think learning is *surprise driven*. Whenever things are predictable or emotionally mild, not much changes in the brain. Why should it change? The world is more or less in our control, we know what’s going on, and there’s no particular reason to change anything. But when we’re surprised, when expectations are violated, and when we encounter novel items—things we’ve never seen before, situations we know we’ve never been in before—then I believe we drop into a different mode in which we note and remember items much more vividly. The more surprising, the more alarming, the more novel, the more is remembered.

Why are surprises remembered? Well, we need to assign credit. When we encounter

something novel, we don’t know what to think of it. We do have to note such items in order to later judge their importance. (Curiosity is also tied with this. We may seek out surprises if they aren’t coming fast enough for us, and our learning rate is stagnating.) We learn disproportionately from intense experiences. Think of your first romantic experience. You probably thought about it for a long time afterward. You probably still remember it pretty vividly, if you’re anything like me. What’s going on?

Consider great accomplishments, things that you worked very hard on and finally achieved, things that were extremely lucky, praise that you’ve gotten from people that you admire, experiences of terror or pain, deaths of loved ones, abuse, humiliation, failures; all these things cause us to experience intensely. In such cases, we note and remember many relatively raw features, both ones that were present at the time that the shocking or surprising event happened and ones from around that time (songs, weather, clothing we were wearing, particular locations, and so on). We tend to need a long time to get over intense events, and we may obsess over them over some period of time. I believe that such obsessing corresponds to substantial reorganizations of our minds/brains.

What’s the purpose of obsessing? Why

...
*the brain and
mind “work
through”
intense
experiences
by repeated
rehearsal,
daydreams,
talking to
oneself. In the
process, we
construct
demons that
can give us
timely alerts*
...

would it be valuable to remember many details, most irrelevant? Imagine that we experience two very similar situations, one case that turns out very positive and another similar situation that turns out extremely negative. How are we to ever learn to predict the difference in the future? If we've generalized the first experience to a high degree, selecting minimal patterns of features, there's no way in which we can later reliably learn the differences between the two, since certain features, now lost, may be essential clues to telling the difference. So I think one important reason we remember such things vividly and go over the features of the experiences is so we can later identify the differences between these situations and differentiate them. We don't necessarily create accurate causal models. We can be haunted by experiences even though we know rationally they may never happen again. Even those of us who are highly scientific may feel at least hints of superstition—"well, things turned out well when I did x , and even though I can't see a causal relation from x to the outcome, I might as well do x again."

What's the value of obsessing? I think the brain and mind "work through" intense experiences by repeated rehearsal, daydreams, talking to oneself. In the process, we construct demons that can give us timely alerts: We can avoid bad things in the future, avert disasters, or we can seize opportunities that present themselves. By rehearsing situations, we can construct demons that give us very early warnings, noting the faintest signs of dangers and opportunities ahead. I think we assign values to people, objects, places, situations, events, and relations by constructing a number of very quick static evaluators. These provide us with an instant impression of whether something is good or bad. There are recent psychological studies by Bargh that strongly suggest that people have quick (often unconscious) positive and negative reactions to just about everything.

Language and mental imagery play a critical role in this. Our basic perceptual systems are fast and feed forward. To modify these systems and build new ones, I think we use internal language and mental imagery to "reexperience" events and "what-if" scenarios, so that we can build new chains for when we encounter similar situations again, as well as hypothetical situations that we have imagined in this process. The results are new, very fast feed-forward perception-action chains as well as self-descriptive prediction systems.

If events are sufficiently negative, they can be debilitating. Battle fatigue and delayed

stress syndrome are examples of experiences so intense that every subsequent experience is likely to contain some reminders of it. That is, certain events can generate so many features—all associated with strong negativity—that any day to day experience will share some features that cause the victim to recollect and relive the negative experience. And I think we all know sad stories of adult maladaptations that can be traced to abuse in childhood.

Animal Intelligence

Much of this is shared with animals. I had hoped to bring a videotape that had been shown by Pat Churchland, but I'll have to settle for describing it to you. We are not alone, I believe, in being smart and self-reflective in exactly the sense I've been talking about. Pat showed a video of a baboon being put in front of a mirror. This was a large mirror, a wall-sized mirror, very clear and sharp. The baboon looked at the mirror, saw another baboon, lunged at "the other baboon," which in turn of course lunged back at her. She jumped back in surprise, and then made a threatening movement, mirrored by the other baboon, and she then just ran off in a panic. The video then showed a chimp being put in front of the same mirror. The chimp had initially the same reactions: She jumped toward the mirror, and the other chimp jumped toward the mirror. The chimp then just stood back, making some faces and hand motions and watching the "other chimp" doing the same things. And then there was a cut to a scene where the chimpanzee was turned back to the mirror, examining her own rump and then up close, looking in her own mouth, with an expression that conveyed "ooh, what's this?" It was really startling. There was no question that the chimp was self-reflective, that she realized that not only was she seeing herself, but also realizing that she had found a new method for discovering things about herself, things she'd never been able to see or know before.

Animal Superstition


Konrad Lorenz was very fond of geese and imprinted himself on many of them. He'd be present when they hatched, and they would imprint on him and follow him everywhere. He was essentially their mother. He had one particular favorite goose who followed him everywhere, except that the goose would never follow him into his house. Geese are said to be afraid of going into dark places, so the goose would always stop at the door. But one day, the

goose absent-mindedly followed Lorenz into the house and suddenly found herself in this completely dark place and panicked. When geese panic because of darkness, they'll fly toward light. So she flew to a window and stood by the window and cooled down, and after her eyes adjusted to the darkness, she resumed following him around the house. Thereafter the goose followed him into the house every day: The goose would follow him in, go and stand by the window for a while, and then return to following him. But gradually, over time, the goose started only making a sort of detour toward the window, no longer standing by the window, but just making a deviation in her path. One day the goose absent-mindedly forgot to make the deviation in path after coming in the door, suddenly realized it, flew in a panic to the window, and cooled out for a while. So what was going on there? Lorenz argued the goose was truly superstitious, that somehow, even though the goose now knew the house well, and felt safe there, she needed to make that little deviation in her path, or else who knew what could happen? We humans aren't so different.

The Bad News

AI has for the most part neglected these sorts of issues. And I think this is a serious problem, because it raises the question about whether AI, if it continues on its current course, is really up to the challenge of fulfilling its stated long term goals. I would argue that most AI research to date has been dictated largely by applications that people have chosen to work on or by available methods. And so it's rather like looking under the light for your keys even though you know you lost them somewhere else because it's easiest to see there. AI has ignored most of neuroscience—some neuroscience knowledge has come into AI via the neural nets community, which has made a serious attempt to follow neuroscience—but AI really hasn't taken neuroscience seriously, even though it is a very fast growing and relevant body of knowledge. It's very clear that fMRI (functional magnetic resonance imaging) studies support society of minds and/or PDP-like models; mental activity is highly distributed. This hasn't had much of an impact on AI. And it ought to because, even as a practical matter, we are in the age of distributed computers.

I'm going to talk about these three topics: (1) perception, situation assessment, and action; (2) meaning and deep language understanding (as opposed to statistical processing or web searching); and (3) generativity and creativity.



... experiments show that young children, too young to have been able to explore the world with their hands, understand in some fairly deep way the physics of the world

Perception, Situation Assessment, and Action

Most of the AI work in perception, situation assessment, and action concerns symbolic reasoning about situations. AI models already assume that somehow we've discovered the objects and relations that are important in a situation. Very little work has been done on the problem of actually turning real, complex scenes into symbolic situation descriptions. And I would argue that this is where most of intelligence really lies.

Perceptual pattern recognition is in a very primitive state. Human Go champions face no threat from AI programs anytime soon. Why is that? Well, people are very good at and, I believe, very dependent on being able to make perceptual pattern judgments. In chess, this is also true. But chess is a sufficiently easier game that has been attacked successfully by computers in a very different fashion.

Why hasn't AI looked at this sort of thing more carefully? I think there are two main reasons: One is that perception really is complex and difficult. It depends on extensive innate wiring. Broad principles just aren't going to work. There are some principles, but there are a lot of them, and few apply generally. Some examples: Animals clearly have extensive innate perceptual abilities. Animals have mating rituals and dances that are extremely elaborate (another topic of Lorenz if you're interested in details). But clearly these are not learned; they're innate in the creature. Ungulates—cattle, sheep, horses—can walk and they can find their mother's udders, they don't bump into things, and they do all this immediately from the time they are born. They don't learn this. It's wired in. Human experiments show that young children, too young to have been able to explore the world with their hands, understand in some fairly deep way the physics of the world, as discussed later. In contrast, most of AI work in vision—which, by the

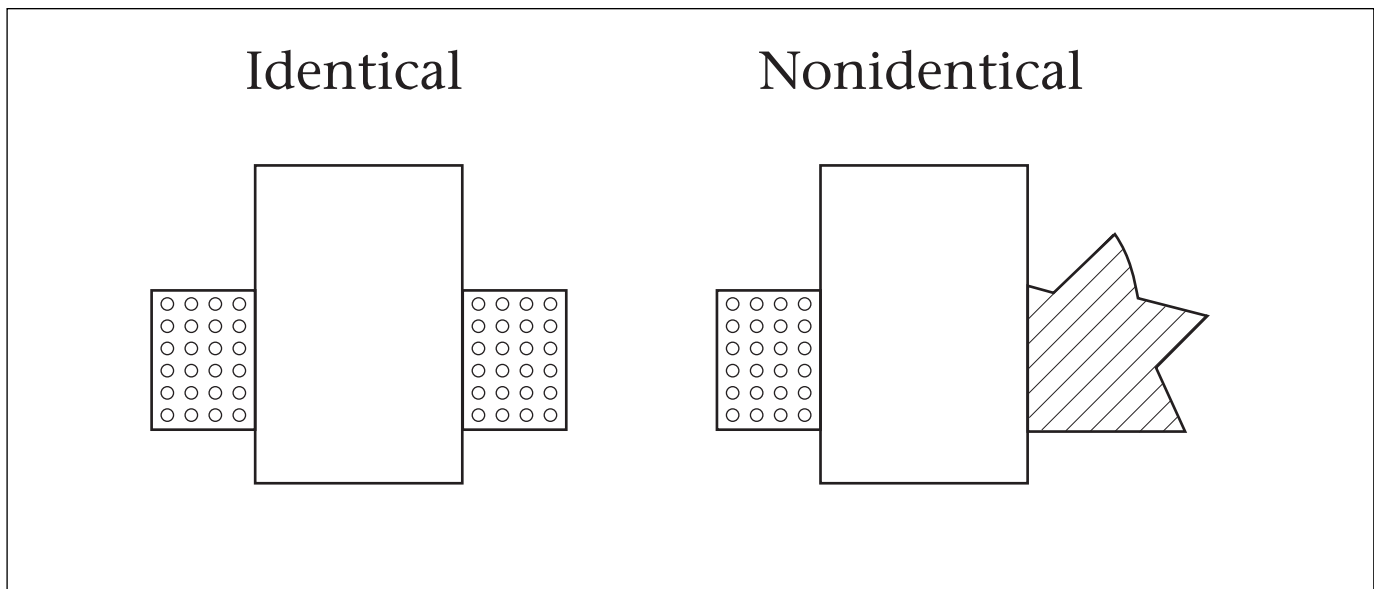


Figure 2. When Two Object Parts Are Seen behind an Occluder, 4-1/2-month-olds Assume There Is One Object If the Parts Are Identical (left) and Two Objects If the Parts Are Dissimilar (right) (Bremnan, Slater, and Butterworth 1997).
 From A. Slater, "Visual Perception and Its Organization in Early Infancy," © 1997. Reprinted by permission of Psychology Press Ltd.

way, I think is a good and thriving field—has concentrated on low level vision—edge finding, segmentation, object recognition, 3-D structure and so on—and hasn't really treated *perception*, that is, the problem of how one selects what's really important from among all the possible segmented objects.

Evidence for Innate Perceptual Knowledge in Humans

To make clear the importance of innate wiring in human perception, here are some experiments from Renee Baillergeon and Elizabeth Spelke. The basic form of their experiments is this: An infant sits in a seat and looks at some sort of display and if the child attends to the display, then the presumption is the child is interested in it, there is something novel about it, something surprising or interesting, but if the child starts getting bored, looking around and fidgeting, then the presumption is the child is not interested anymore because the displayed situation is boringly familiar. This sounds shaky, but researchers have been able to replicate very strong, clear effects. In the first experiment (figure 2), a 4-1/2 month old infant is shown a scene consisting of three regions, one occluding the other two. In this case, if you show the child the figure where the occluded regions are similar and then pull on one of the regions, but only one piece comes out, the child is very surprised. The child expects that both regions are part of the same

object and that they will move together. If you take a scene with dissimilar occluded regions, pull on one, and both pieces move together, the child is surprised. The child expects because the regions are different that they are not part of the same object. And this is judged by the child long before the child could have figured out such things by playing with objects. Figure 3 shows a similar display: If you show a block with occluded regions moving in a coordinated manner behind it until the child gets bored, you then remove the occluding object to show either a complete rod or two short rods joined invisibly. Now, to a 4-1/2 month old, the left bottom display (the complete rod) is extremely boring: The child has already figured out that the rod is continued behind the occluding object. But the situation on the right is very interesting: The child is very surprised that these aren't really joined. It turns out that before 4-1/2 months, the effect is exactly opposite, and nobody really quite understands why that is. But it can't have been judged from experience because an infant of this age still lacks hand-eye coordination and can't have had any physical experience with such situations.

Figure 4 requires some explanation. The bottom line of the figure depicts a block—the shaded object on the left—and a thin board, which is the object on the right. The child sees the board rolling up to and then past vertical, coming to lean on the block, then coming forward and lying flat on the table. And that's

dull and uninteresting to the child. The middle display, an impossible event, represents the following: The board rolls up, goes back to where the block should be but is not supported by the block (which has been removed by the researcher), allowing the board to go flat on the table. The board then rolls forward and then the block appears; so it's as though the block somehow disappeared and reappeared during the process. This is very surprising to infants. So in some sense they really understand the object ought to be there, that it ought to support this board, and if it doesn't, something is really wrong. So there has to be a lot of innate wiring—really a lot of knowledge—that's built into us. To build a perception system, that information somehow has to be there. It's not likely to be learned easily—it was learned evolutionarily, but for AI programs, it would have to be programmed in.

Links between Perception and Action

The second reason that perception is difficult is that perception is goal-directed and, therefore, varies as goals change: Perception is tightly linked with a rich set of goals and actions. Of course, not all aspects of perception are goal-directed. Some parts of scenes are inherently salient and “stand out”: things that are very big, or moving very fast, or very bright, or familiar (for example, people you know). But in general, perception is responsible for assessing scenes relative to the assessor's goals and experiences. You see in the scene opportunities and dangers, things you can eat, things that can eat you, metaphorically of course. There's a large body of work by J. J. Gibson from the 1950s and 1960s that talks about perception in exactly this form. Gibson described the perceived world as consisting of *affordances*, for example, the things that the scenes offer to us, the possibilities and risks. I don't think general purpose representation is even possible, except in the very, very simplest domains, where one has, for instance, plane surfaces with polyhedra. How do you make sense of complex scenes and situations?

To illustrate this in part, consider figure 5. This is a picture from the *National Geographic*, and at the first level of analysis, it's an orangutan in a tree. This is a hard scene to parse with many very small regions. I'd like you to imagine yourself in the following situation: Suppose you're really in the scene, and this is what you're seeing, you're there. If you're a photographer looking at that scene, what would you see? I would argue that first of all

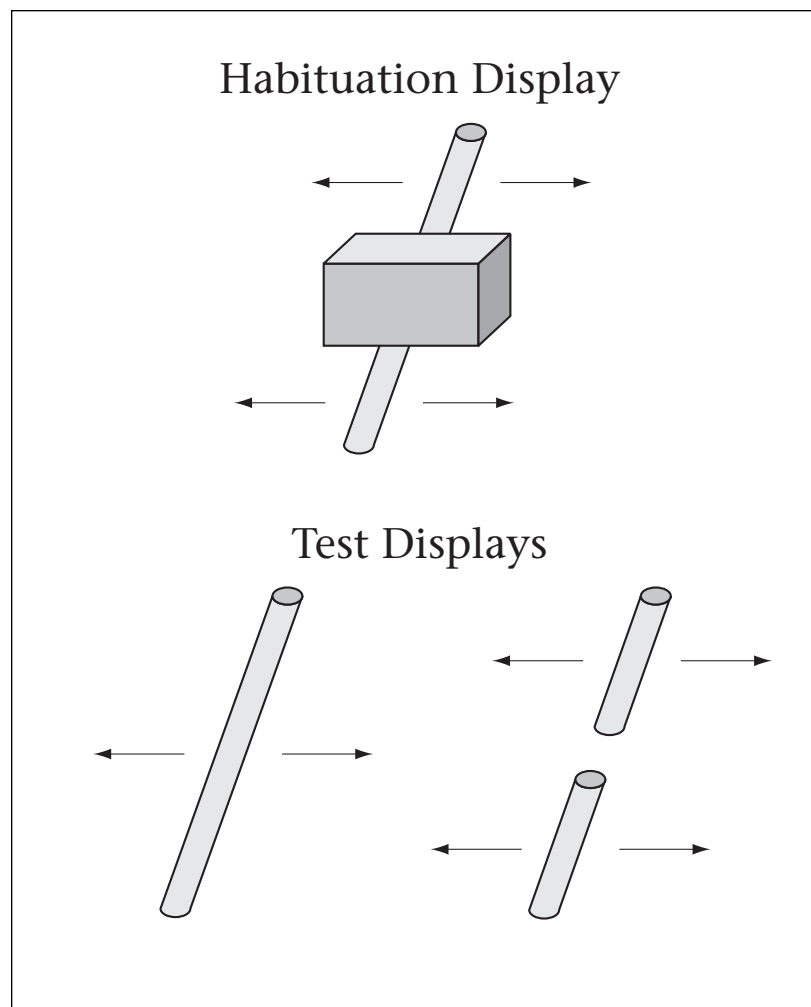


Figure 3. Habituation and Test Displays in Experiments' Perception of Partly Occluded Objects.

During habituation, the rod moved back and forth behind the occluder (Bremnan, Slater, and Butterworth 1997). From A. Slater, “Visual Perception and Its Organization in Early Infancy,” © 1997. Reprinted by permission of Psychology Press Ltd.

as a photographer for *National Geographic*, when you look at this you'd want to be able to show enough of the surroundings to give an idea of the habitat and terrain in which this orangutan lives, how much visibility there is, how much she wants to be hidden versus being able to see what's coming in terms of predators, what kind of foliage or vegetation this creature might eat, whether she's in a nest or not. You'd want to show how high off the ground she is, have a pleasing composition to the picture, and so on. Now suppose instead that you are a collector of animals for a zoo, then what would you see? Well, I think you'd see something very different. You first of all might look and say, “Hmm, can I sneak up on this creature? Is there any way for this creature to get away?”

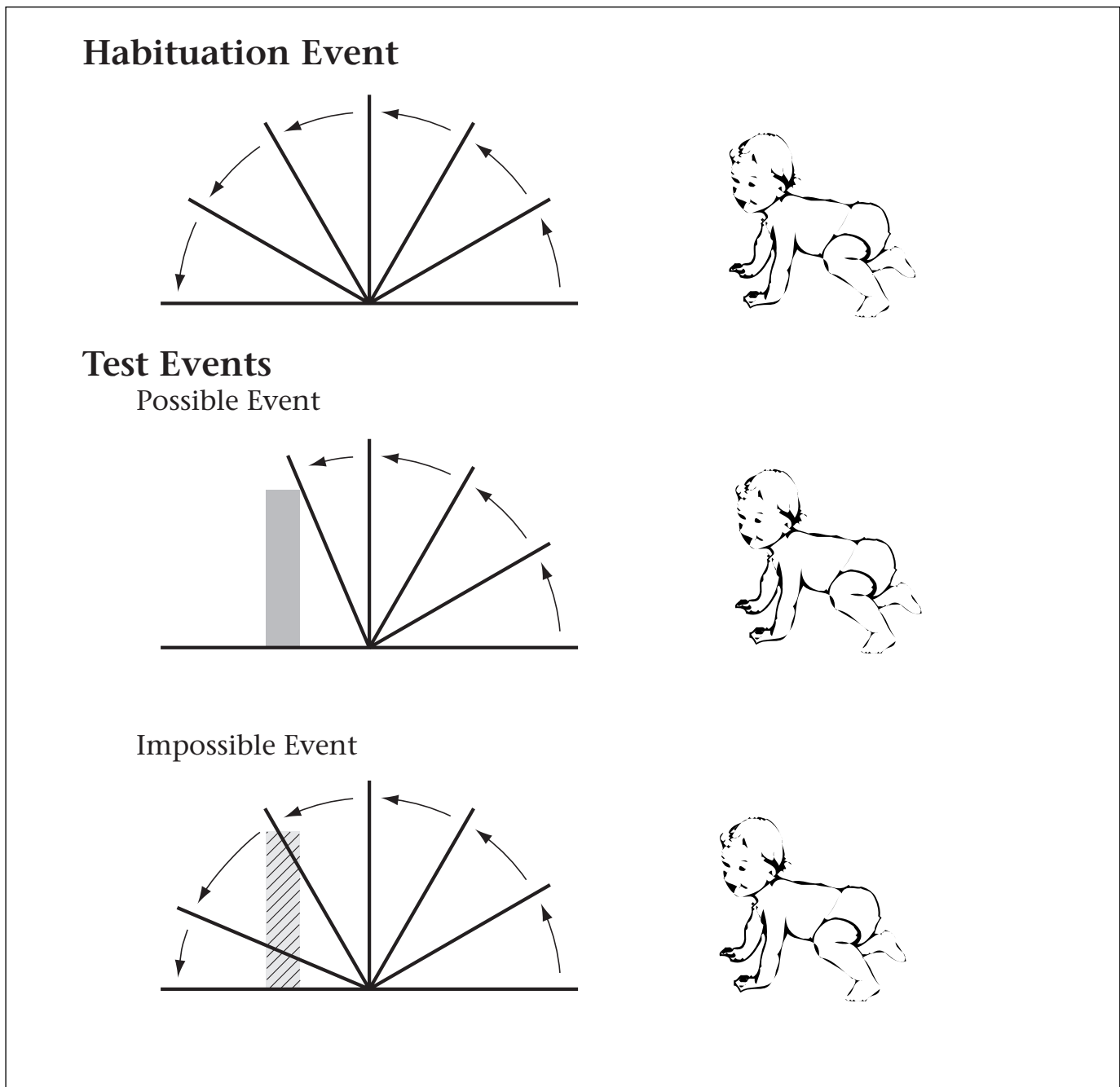


Figure 4. Schematic Representation of the Possible and Impossible Test Events Used in the Principal Experiment (Baillargeon, Spelke, and Wasserman 1995) (courtesy of Renee Baillargeon, University of Illinois at Urbana-Champaign).

Could the creature jump to this nearest tree? Could I get there in time? Is there a path? How would I put something, a net or something, out to catch it? Could I hit her with a dart from here?" You'd be looking at very different parts of the scene and in different ways. Third, imagine that you are a member of the World Wildlife Federation, knowing that there might be a collector, a poacher, nearby. Then what

would you look at? Well, I think you'd look around the bases of the trees and ask yourself, "Is someone sneaking up on this creature. Is there some way I could intercept such a person in time, etc.?" You'd look at different parts of the scene, you'd look at them in different ways, and you'd look at them in terms of paths, trajectories, distances, escape routes, accessibility routes, and so on.

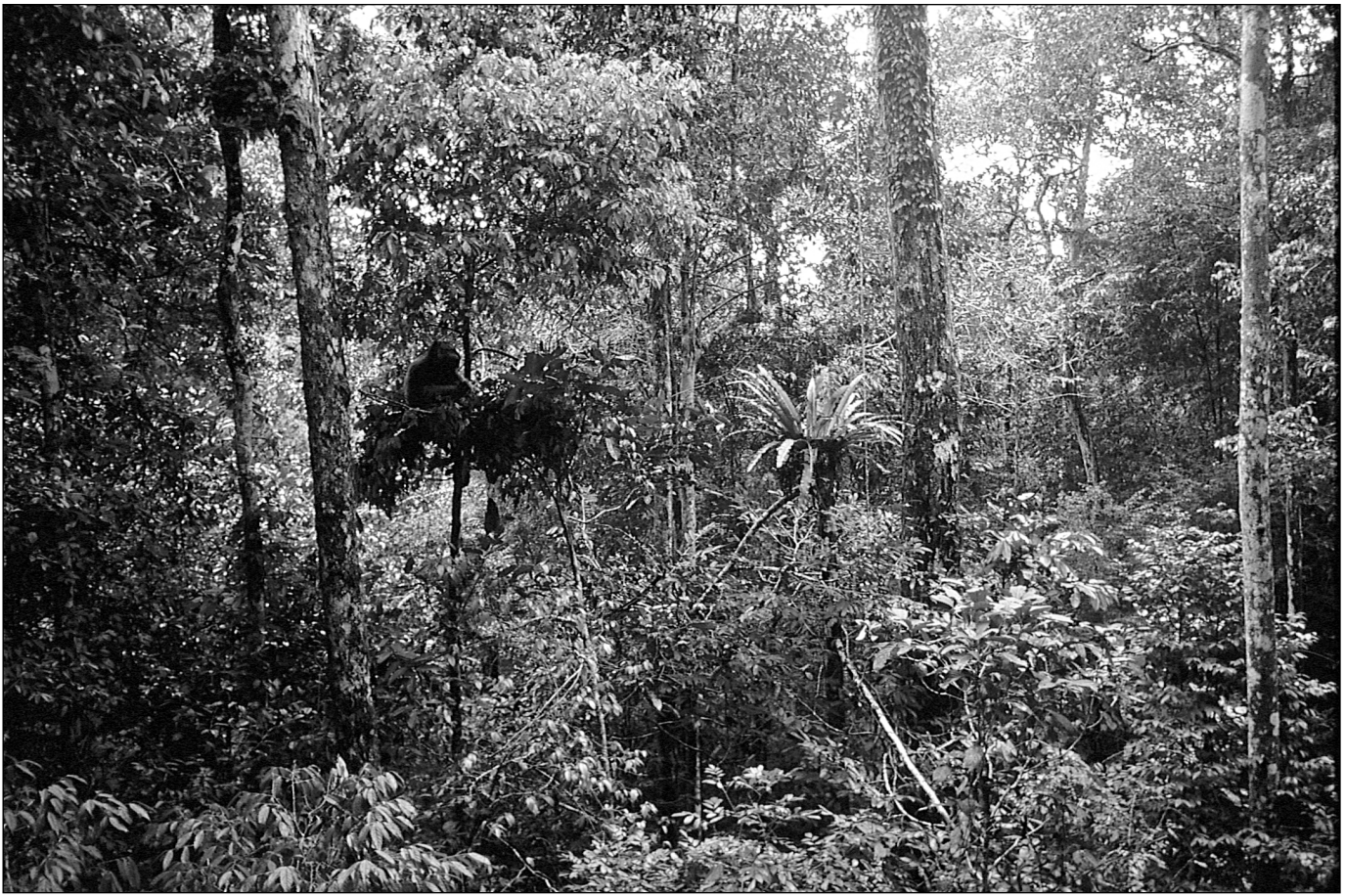


Figure 5. *A Complex Scene* (Knott, C., and Laman, T. 1998. *Orangutans in the Wild*. National Geographic 194(2): 30–56).
 Courtesy Timothy Laman, National Geographic Society Image Collection.

AI Successes in Perception and Language

Now, I don't want you to think that all the news is bad. There are a number of AI successes and promising projects that address the issues I've been discussing. I'll list just a few. Successful systems that have perception and action for the most part don't do anything terribly complex, but they do take real world scenes and do something useful with them. I mentioned RALPH, the autonomous driving system used in the NO HANDS ACROSS AMERICA project. RALPH looks at the real world, and it finds travel lanes, other vehicles, road obstacles, hazards, exits (to avoid), intersections, and so on, and it does something with them, namely, it changes the direction of the car and may hit the brakes. Autonomous planetary explorers like SOJOURNER have been pretty dumb to date, but SOJOURNER's descendants will do things such as locating interesting targets, for example, rocks or their home base, devising paths to get to these targets while avoiding obstacles such as big rocks

or cliffs that they might not be able to climb over or might fall down into. And there is a large body of work by Rod Brooks et al. on various reactive robots that interact with the real world.

Evolution of Language

Language and meaning are quintessential features of human intelligence. Language is not just for I/O. Early on in AI's history there were influential articles that argued that the central core topics of AI were logic, reasoning, and knowledge representation and that language and vision were peripheral operations for the central reasoning core. I disagree; language is central, not peripheral.

The origins of language are unclear, but it's fun to speculate. My guess is that language is the intersection of (1) the ability to reify objects perceptually, that is, the ability to create coherent mental chunks that can be named; (2) the richer association of possibilities that a larger brain offers—more possible

connections; (3) metaoperator development that lets us learn via self-observation and rehearsal as I mentioned earlier; and (4) preexisting signaling systems, shared with animals. I believe there may have been a synergy between these abilities that led to a critical advantage for early ancestors and eventually developed quite significantly, giving huge advantages to hunters and gatherers and ultimately leading to civilization. (There are many other proposals. One seriously advanced theory is that gossip is the root of language; that gossip is extremely important for social bonding, stratification of society; and consequently, that we are wired to love gossip, much as we're wired to love food and sex. It seems to be true that most of us love to hear and tell juicy stories about the people around us. Language thus developed initially to let us gossip more successfully! Others have suggested that language developed because women preferred poetic men. I think that these factors may have played some role, but not the central one.)

I think language evolved through interlocking synergies. The relatively undeveloped state of human newborns made it possible for us to learn a great deal more, including language, because we were born incompletely developed. There's a natural course of innate development that would occur anyway, but its direction and content are influenced by outside experience—by the language heard around us, by perception, by the particular circumstances we find ourselves in. In turn, of course, the helplessness of lengthened childhoods requires stable families and societies. And so language itself may also have played a role of increasing the fitness of families and societies that made possible the longer childhoods and support of children until they were able to go off on their own. And so these forces may have formed a synergy that was important in language development. Language makes culture possible: we don't even have culture without language. *Memes*—the things that we name and share as the standard units into which we divide the world—are language items. They might be things that we discover ourselves perceptually, but they might not be. Language makes them inescapable. And of course language itself is a major element of culture. Language itself can substitute for direct experience so much of what I said about perception is also true of language. Language allows us to experience things that we can't see directly, things that are far away in space or time, as well as accumulated culture, and more abstract things: morals, hints, metaconcepts, stories, hypotheticals, rules, and so on. Language is self-extending, so

that once a certain level of proficiency is developed, people can reify objects and events even if they haven't seen them perceptually. (Perhaps we can't even avoid reifying experienced items.) And I think we need importance judgment for abstracting correctly—there are a lot of ways we can abstract. How do we do it properly? And we clearly need judgment for understanding things such as anaphora, humor, fables, or parables. A fable isn't just a story about some creatures. It's really about some bigger issue, and we somehow learn early to understand that.

Language is related to and depends on perception. The possible meaning structures for utterances are more constrained than in the case of perception, which is much more open-ended. But many of the arguments made for perception also hold for language. Language can have special meanings and affordances for the understander: You can hurt or inspire somebody very effectively with language. Probably for many of you the most intense experiences you've had—positive or negative—have been language-mediated rather than perceptually mediated. Some of what your parents said to you, your boss said to you, your child said to you have likely had a huge impact.

Much language requires perceptual reasoning. I wrote an article a long time ago when I was already thinking about these sorts of things. The central example I used was, "My dachshund bit our postman on the ear." Now why is that an odd sentence? It's odd because postmen's ears aren't ordinarily anywhere near dachshunds' mouths. So one needs to invent some sort of story for how this could happen, how this could possibly make sense: maybe the postman fell down, or the postman picked up the dachshund to pet him. But this requires perceptual reasoning, possibly involving the perceptual system, even though the similar sentence, "My doberman bit our postman on the ear," is not problematic—a doberman is big enough to reach the ear. But the fact that the perceptual system can be pulled into action if necessary, is, I think, significant.

How can we build systems that understand language? I think text based learning methods, statistical methods, simply won't do it all, unless we somehow solve perception and can use perception to support language learning from experience, which is a big if. Humans learn language with the aid of perception. Symbol grounding is very important. Language experience has to be, to be really understood, linked into the structure of the corresponding situation, which, like language, also

Capture	Set First	Play	Save	Set Las
Beginning	First 30	-T 30	+T 30	Last 50
Jpeg	Mpeg1	640x480	320x240	-sec 3
Viewfinder	Images	Motion	Objects	Ellipse
Red	Green	Blue	Yellow	
Uninterlace	Deinterlace	Corresponded	Autoclip	Classif
-sequels 3	+sequels 3	-sequel 0	+sequel 0	
Capture All	Autoclip All	Save All	Train	

Figure 6. Video Sequence to Be Classified as an Event by HOWARD (Siskind and Morris 1996) (courtesy of NEC Research Institute).

has meaning and context. As with perception, I believe that much of language is innate. Chomsky's main point is that syntax is largely innate, and I think this is moderately well established, although the scope of innate functionality is certainly still arguable. Less arguable are the facts that deaf children spontaneously sign and that twins are frequently observed to invent private languages of their own. And there are many other examples.

I'm going to show you a video of HOWARD, one of the best systems for actually understanding real, complicated, moving scenes. This is work by Jeff Siskind who's at NEC Research Institute now. The work was originally done while he was at the University of Toronto. Figure 6 shows a sample from Jeff's current follow-up of this earlier research. The following is the audio track of the videotape:

People can describe what they see.
They describe not only objects like blocks,

but also events like picking things up. This video presents Howard, a computer program with the ability to describe events. Our approach is motivated by a simple observation: We pass movies of events through an edge detector. While people couldn't recognize the objects from stationary edges alone, they could recognize events depicted by edges in motion.

An event pass can be described by the motion profile of its participant objects. A picking up event has two sub-events: First, the agent moves toward the patient while the patient is at rest above the source. Then the agent moves with the patient away from the source. Our event models attempts to capture these motion profiles.

We don't use detailed image understanding to perform event recognition.

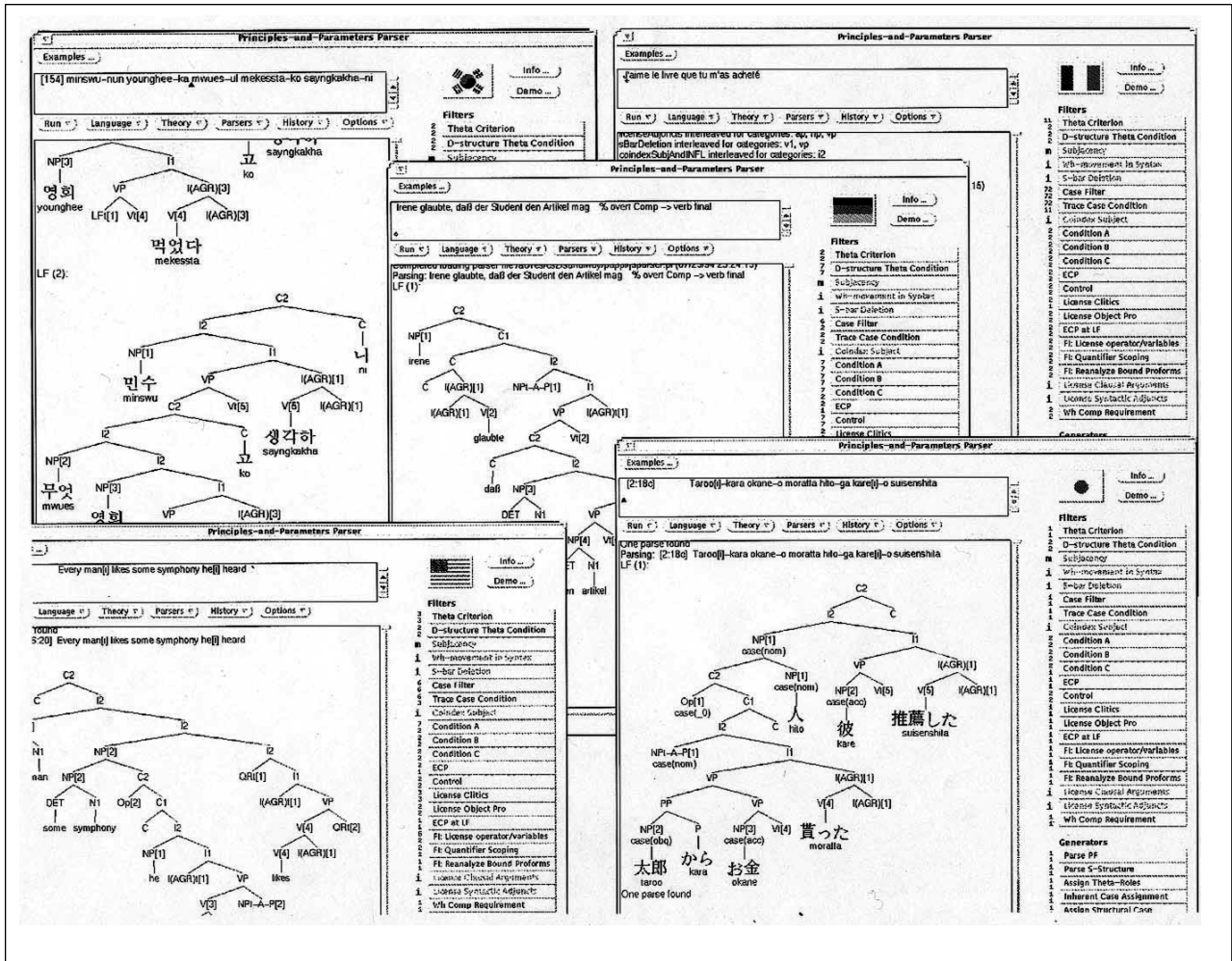


Figure 7. Parse Trees for Sentences in English, Korean, and German (courtesy of NEC Research Institute).

Only the rough position, orientation, shape, and size of the participant objects are needed. We represent this information by ellipses centered on the participants. Our training and classification procedures operate solely on the streams of ellipse parameters produced by our tracking procedure.

We process the movie one frame at a time to find regions of pixels that are either moving or are brightly colored. First, we ignore pixels with either low saturation or value. Then we group nearby similarly colored pixels into regions discarding those that are too small. Optical flow is then used to find moving objects in the non-brightly colored pixels. Next, we group nearby moving pixels into regions, again discarding those that are too small. We use both the colored and the moving regions to track the object. Finally, we fit an ellipse to each region.

Subsequent processing uses only ellipse data.

Here you see the colored and moving regions found by applying our tracker to a sample movie. Once the ellipses are placed in each frame, we find the ellipse correspondences between frames. Our technique attempts to fill in missing ellipses and filter out spurious ellipses.

We processed 72 movies with our tracker. Of these, 36 were randomly selected as training movies, 6 movies for each of the 6 event classes. We constructed a single hidden Markov model for each event class. We then classified all 72 movies, both the original training movies and those not used for training, against all 6 event models. Our models correctly classified all 36 training movies and 35 out of the 36 test movies.

You are watching the results of our classifier now. Different events can have dif-

ferent numbers of participant objects. Our pick up and put down models have three participants, while our push, pull, drop, and throw models have two. Movies are classified against only those models with the same or fewer numbers of participants as the ellipses found by our tracker. Our classifier can correctly match a subset of the ellipses in this three ellipse movie to the drop model, which has only two participants. It appears that poor tracking caused one drop movie to be misclassified as a throwing event.

I should comment that Jeff can now show you a real time demo on his desk in normal light thanks to computer power increases. He's trying to significantly increase the set of possible events. On one hand, that doesn't seem very impressive: It's a long way from being able to do what you'd really like a system to do, namely, describe what's going on, given any possible scene. On the other hand, the real world, with its shadows and complexity, is hard to work with, as any of you who have tried it can attest.

Overall, to really build intelligent systems, I suspect that large amounts of handcoding are going to be essential. Fortunately, a lot of work has been done in recent years. We have examples such as WORDNET and CYC. In the subsections to follow I'll tell you about the PAPPI system and work by Inquizit, Inc., that I recently encountered. Inquizit has built a huge natural language processing system with many levels and functionalities.

PAPPI

PAPPI is the work of Sandiway Fong of the NEC Research Institute. PAPPI parses a broad range of difficult sentences from 12 languages. It's a "principles and parameters" parser, inspired by Noam Chomsky's ideas. Figure 7 shows parse trees for sentences in several languages: English, Korean, and German. PAPPI comprises about a quarter million lines of PROLOG code, and it's able to handle new languages by simply setting parameters correctly, an impressive feat. Sandiway's group has been able to add new languages, including complicated languages such as Hungarian or Turkish, based on one summer's work with a linguistics grad student for each language.

INQUIZIT

INQUIZIT, a product of Inquizit Technologies, is a very interesting system, written by Kathy Dahlgren and Ed Stabler over a 14 year period. INQUIZIT has an extensive word sense lexicon with 350,000 word senses as well as a morpho-

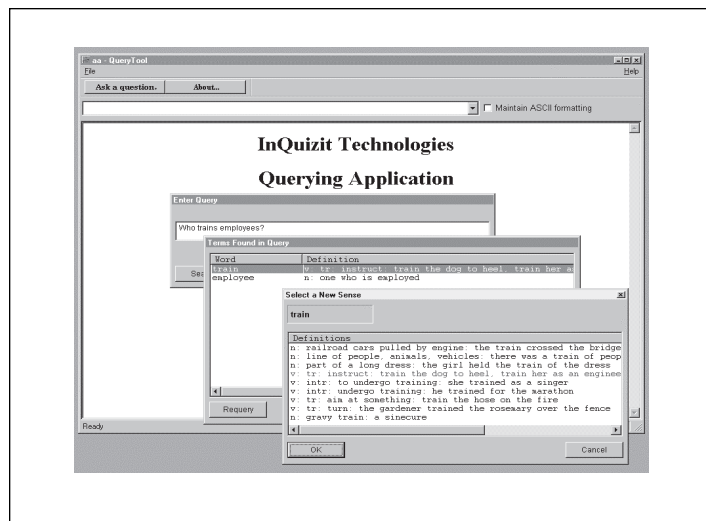


Figure 8. INQUIZIT, as Applied to an Employee Manual, Answers the Question "Who Can Train Employees?" (courtesy of Kathy Dahlgren, InQuizit Technologies, Santa Monica, California).

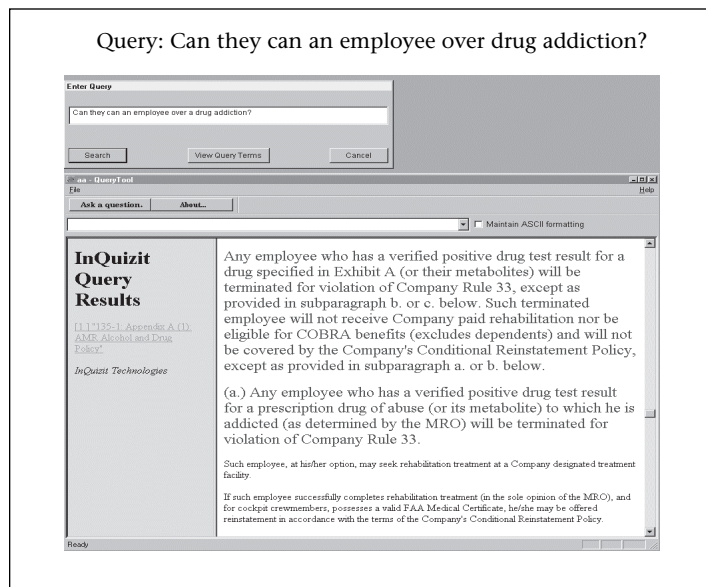


Figure 9. INQUIZIT with the Question "Can They Can an Employee over Drug Addiction?" (courtesy of Kathy Dahlgren, InQuizit Technologies, Santa Monica, California).

logical analyzer that lets it actually represent over a million word senses. INQUIZIT has an ontology with inheritance, and a "naive semantics" with first order predicate calculus tests to determine when the various definitions can apply. Figure 8 shows INQUIZIT applied to an employee manual, answering the question, "Who trains employees?" The point here is to show that INQUIZIT knows a large number

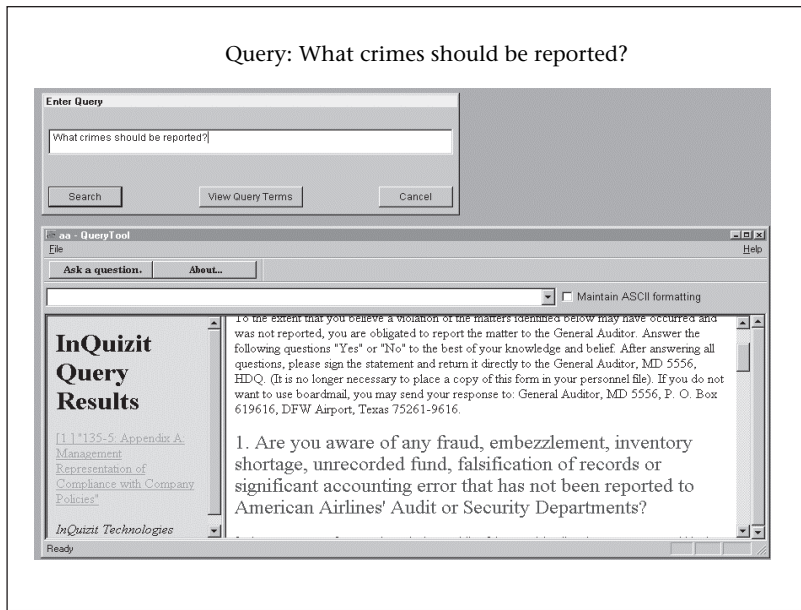


Figure 10. INQUIZIT with the Question “What Crimes Should be Reported?” (courtesy of Kathy Dahlgren, InQuizit Technologies, Santa Monica, California).

of meanings for “trains”—railroad cars, a parallel line of people, and so on—and it’s able to pick the correct meaning—“instruct”—based on the values of predicates it applies to the surrounding sentence. Figure 9 shows a sentence being used to do text retrieval; the highlighted text is what’s retrieved. The question here is, “Can they can an employee over drug addiction?” The example illustrates first that INQUIZIT can pick the right meanings for each occurrence of *can*, but more importantly that INQUIZIT can match text based on deep meanings—note that the word *can* is never used in the matched passage. INQUIZIT understands that in this case *can* means *terminate*. Figure 10 shows a third example, “What crimes should be reported?” In this case, the word *crimes* never appears either but the text matched says, “Are you aware of any fraud, embezzlement, inventory shortage?” and so on, and INQUIZIT is able to use its ontology to judge that those words are specific examples of crimes and match them. Building PAPP1 and INQUIZIT (and CYC) has required a huge amount of hand effort. The bad news is that I think we will need to tackle a number of similarly large tasks in order to build systems that are really intelligent. The good news is that it’s possible that this can be done successfully for very large—but finite—domains.

Generativity and Creativity

Art, science, and technology depend critically on our ability to create ideas and artifacts that have never been seen before and to create novel relationships between people, objects, and ideas. Doing this depends clearly on developing meaning in perceptual pattern systems and on creativity, but creativity is not just the province of arts and sciences. Infants and adults of all ages illustrate creativity: infants by their play activity, by their language, by understanding, and use of metaphor and analogy. We aren’t going to have systems we ever regard as smart if all they do is react to what they see. And in fact they can’t really understand language unless they are able to make some sense of analogy and so on. And they won’t be able to express things very concisely unless they can also use language and use it appropriately. I know this is important, but I’m not certain how to attack this as an AI problem. Some reasonable starting places can be found in the work of Ken Forbus and Dedre Gentner, Maggie Boden, Harold Cohen, John Koza, Doug Lenat, and Doug Hofstadter.

The Role of Logic

A possible objection to what I have said so far is, “Isn’t everything built on logic?” A substantial community within AI has long held the view that the human brain somehow implements a perfect logic machine and that all intelligence is built on top of this. A corollary to this view is that in order to create a fully intelligent system all we need to do is create a perfect logic machine and then build everything on top of it. I’d ask you the following question: If that’s true, why are adults so miserable at reasoning and hopeless at formal reasoning for the most part? (None of you of course. You’re all great at it! But if you take the overall population, they’re not so great.) We have cases of people with Williams syndrome—kids who speak extremely articulately, who have great understanding about themselves, social situations, and other people, but who are absolutely hopeless in reasoning about the physical world, and have very low IQs. If intelligence was based on a common logic reasoning system, how could such a case exist of people who are so good in language and social reasoning but so poor about most other things?

My view is that logic is really the fruit of learning, one of the **last** things we learn, not the system on which everything is based. Logic is the culmination of repeatedly detecting pat-

terns and abstracting them from specific events and schemas and then from these ever more abstract schemas. However, I think that when we do this, we also retain polymorphic features so that we can apply the patterns and schemas if and only if they are appropriate. As evidence, we have work by Tversky who showed that people are much better at logic problems when they're stated as social situations, but when exactly the same schematic problem is stated as *As*, *Bs*, *xs*, and *ys*, people are miserable at solving them. We understand envy, jealousy, competition, and such stuff well, but we have trouble reasoning about sets and relations. We also have such gems as the competing proverbs: "Absence makes the heart grow fonder" and "Out of sight, out of mind." They're both sort of right, sometimes. We know when one or the other applies.

Logic lets us talk about unimportant things. It lets us take diverse objects—important and unimportant—and turn them into objects of equal importance. In logic, all objects, operators, and so on tend to be uniform, and this can be extremely important and useful, but it's very different I think from what intelligence is usually doing, namely, judging what's important under the circumstances.

Summary

Perception, action, language, and creativity are central to intelligence. They aren't just I/O functions, they aren't just peripheral operations that AI's going to need to connect the **real** systems to the outside world. There's very little—perhaps no—evidence for a central CPU. There's lots of evidence for the "society of mind." What humans think is important really ought to have a role in AI—how could it not? Understanding systems need goals and values to guide their actions. If they're to act appropriately, they need to make importance judgments and choices. And of course if intelligent systems are going to understand us, they'll need to be able to model us. They're going to need to be able to understand why we've done what we've done, why we say what we say. If they can't do that, they're going to be pretty useless as colleagues or assistants. So even if they didn't need to have their own emotions and goals, they'd have to be able to simulate them in order to deal with us.

So what should AI be doing? I certainly don't want you to go away thinking that we should stop doing applications. I am also very much in favor of continuing research on reasoning, logic, and so on—it's great work. Applications are absolutely critical for contin-

ued credibility and funding. The work on logic, reasoning, and knowledge representation has been the heart and soul of what lets us do the AI applications we have done. Don't stop! But as a field we need to do more. We need to have a greater emphasis on science. We need to be more problem driven, not solution driven. Even if we have great hammers, we have to avoid the tendency to see the world as made up completely of nails. I deeply believe we need broader scholarship. To do AI properly, we have to know a lot of areas. We have to know about neuroscience, we have to know about psychology, we have to know about linguistics. This is hard. There's a lot to know. I strongly recommend that you cultivate relationships with people who can be good informants so you don't necessarily have to study the whole literature. When you discover parts of the literature that you really need to know, go and learn them. We aren't going to accomplish our largest goals unless we are prepared to do that. We need this both for inspiration and to avoid reinventing the wheel—or reinventing it hexagonally. Future AI applications will really require us to model language, perception, and reasoning well. Other people can use statistics, other people can use simple stuff that we've already shown how to do. What makes us special? We need to have something new. We need to have something that's better.

Service

That's the main thrust of what I wanted to say technically. I do want to add a postscript. This is very important. In addition to research, I really believe that AI has to give much more attention to service. This was underlined in a recent trip that several of us made to NSF, the first of a series of trips I hope we will make to other agencies as well, part of an education process. The story we heard over and over again was that AI has not given sufficient attention to volunteering (nor has CS in general). AI is not well understood by other people in CS, not understood well by funders, by legislators, or by the public, and it's not going to be well understood unless people in AI are willing to go and be part of the decision making and contracting functions of government in our country—and I'm sure this is true of other countries as well. This is a critical step toward influencing policies and funding.

Service is also critical for conferences, for editing of journals and conference proceedings, and prompt reviewing. Review and return your papers more quickly. Let's not make excuses for why it takes two or three years for papers to appear. We can do better

Source List

Creativity

Gentner, D.; Brem, S.; Ferguson, R.; Markman, A.; Levidow, B.; Wolff, P.; and Forbus, K. 1997. Analogical Reasoning and Conceptual Change: A Case Study of Johannes Kepler. *The Journal of the Learning Sciences* 6(1): 3–40.

RALPH

Jochem, T., and Pomerleau, D. 1996. Life in the Fast Lane. *AI Magazine* 17(2): 11–50.

Chess

IBM. 1996. DEEP BLUE Chess Program. IBM Research Products and Projects. Available at www.research.ibm.com/research/systems.html#chess.

General

Waltz, D. L. 1997. Artificial Intelligence: Realizing the Ultimate Promises of Computing. *AI Magazine* 18(3): 49–52. Also available at www.cs.washington.edu/homes/lazowska/cra/ai.html.

Dachshund-Postman Example

Waltz, D. L. 1981. Toward a Detailed Model of Processing for Language Describing the Physical World. In Proceedings of the Seventh International Joint Conference on Artificial Intelligence, 1–6. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

SOJOURNER

MarsMap—Demos. SOJOURNER. ca. 1999. Available at img.arc.nasa.gov/Pathfinder/marsmap2/www_demo.html.

Lyme Disease

Fraser, C. M.; Casjens, S.; Huang, W. M.; Sutton, G. G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K. A.; Dodson, R.; Hickey, E. K.; Gwinn, M.; Dougherty, B.; Tomb, J. F.; Fleischmann, R. D.; Richardson, D.; Peterson, J.; Kerlavage, A. R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; van Vugt, R.; Palmer, N.; Adams, M. D.; Gocayne, J.; Venter, J. C.; et al. 1997. Genomic Sequence of a Lyme Disease Spirochaete, *Borrelia burgdorferi*. *Nature* 390(6660): 553, 555. Available at http://bric.postech.ac.kr/review/genetics_molbiol/97/12_11.html.

Syphilis

UPI. 1998. Key to Syphilis Vaccine May Be Found. UPI Science News, 16 July. Available at www.medserv.dk/health/98/07/17/story03.htm.

Star and Galaxy Classification

DIMACS. 1998. Astrophysics and Algorithms: A Discrete Math and Computer Science Workshop on Massive Astronomical Data Sets, 6–8 May, Princeton, N.J. Available at

dimacs.rutgers.edu/Workshops/Astro/abstracts.html.

Animal Behavior

Glass, J. 1998. *The Animal within Us: Lessons about Life from Our Animal Ancestors*. Corona del Mar, Calif.: Donington Press.

Good general reference that generally agrees with my views, but adds significant new ideas on reasoning, our inner monologues, territoriality, and other important topics.

Lorenz, K. 1966. *On Aggression*. New York: Harcourt, Brace.

Perception

Goleman, D. 1995. Brain May Tag All Perceptions with a Value. *New York Times*, 8 August, C1 (Science Section).

Reports on values attached to perceptions, a report on the work of Jonathan Bargh.

Affordances

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Mahwah, N.J.: Lawrence Erlbaum. Available at www.sfc.keio.ac.jp/~masanao/affordance/home.html

Theories of the Mind

Moyers, B. 1988. Interview with Patricia Churchland on the PBS Program “A World of Ideas.”

A brief synopsis of this videotape can be found at www.dolphininstitute.org/isc/text/rs_brain.htm.

Animal Morality and Superstition

Lorenz, K. 1991. Objective Morality. *Journal of Social and Biological Structures* 14(4): 455–471. Also available at www.percep.demon.co.uk/morality.htm.

Robots

AMRL. 1994. Project 3: Motion Control for Space Robotics: Behavior-Based 3D Motion. Autonomous Mobile Robotics Lab, University of Maryland, College Park, Maryland.

This site contains reactive robot design information.

Cañamero, D. 1996. Research Interests, AI Lab, Massachusetts Institute of Technology. Available at alpha-bits.ai.mit.edu/people/lola/research.html.

General information, including the Zoo Project, headed by Rodney Brooks, is presented.

Moravec, H. 1988. *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, Mass.: Harvard University Press.

than that. Treat reviewing as an important activity, not a nuisance that you only get to after you’ve done all the more important stuff.

Executive Council Actions

The last thing I’d like to do is to tell you about what the executive council is doing. I tried to

decide when I became AAAI President what was the most important thing to do. I decided that the most important thing to do is to make some real use of the large amount of money that we’re fortunate enough to have as an organization, to help AI become a better field—not to merely save the money for a

rainy day. I think we should be careful that it's not wasted, but I think we really can use the funds we have to do something great.

Here are some of the actions that the council has agreed to recently:

We're going to create a number of new prizes for AI members. I think this is important because nobody gets a Nobel Prize (or Turing Award) without having gotten other prizes first. In fact you're unlikely to get a prize from any other computer society unless you've already gotten one from AI. We need to recognize the great work that's done in the field and need to create heroes and heroines and help people's careers.

We're going to create high school and undergraduate prizes to encourage people so they will want to go into AI as a field and so that they will know what AI's about.

We'd like to sponsor two international science fair teams to come to AAAI, find ones that are doing great work that happens to bear on our topic and perhaps encourage them.

We want to put some AI technology on our web site. Our web site ought to somehow reflect the great things that are going on in the field. It ought to be distinctive. It shouldn't be lagging behind the rest of the field.

We're going to try to commission a science writer or perhaps several science writers to write about AI. If the word isn't getting out the way it should be, let's make sure that we hire someone who's persuasive and articulate and tell them what we're doing and let them make our story.

To encourage people to go into government service from within the field, we're going to create service awards for AI people who serve in government positions. Bruce Buchanan has already done a lot of work on creating high school information web pages and brochures that should appear fairly shortly. We get a lot of requests from high schools, and we're going to try to answer them better than we have in the past.

Finally, we're going to increase the budget for the national conference scholarships to let more students, who could otherwise not afford to come, come to the conference. Next year we're going to try something interesting called "CHI-Care." This was Jim Hendler's suggestion. The CHI conference has offered this for a while now. Basically, they have children

take part in the conference program: The kids write a newsletter about the conference, take pictures of the speakers, interview them, mingle with the conference participants, help out at booths, whatever. Apparently this is hugely successful—a lot of fun for the kids and also another way of getting them interested in the field. Next year we're going to leave the conference fee the same, but we're going to include tutorials, so everyone can go to them. We're going to try to continue to increase the cooperation with collocated conferences and continue the series of educational visits to NSF, DOD, NLM, and NIH funders.

And finally we're going to take a member's survey and act on the good ideas.

Thank you very much for your attention. I actually managed to finish before the end of my hour. Have a great conference!

References

- Baillargeon, R.; Spelke, E.; and Wasserman, S. 1995. Object Permanence in Five-Month-Old Infants. *Cognition* 20:191–208.
- Bremner, G.; Slater, A.; and Butterworth, G. 1997. *Infant Development: Recent Advances*. Philadelphia: Psychology Press.
- Knott, C., and Laman, T. 1998. Orangutans in the Wild. *National Geographic* 194(2): 30–56.
- Siskind, J. M., and Morris, Q. 1996. A Maximum-Likelihood Approach to Visual Event Classification. In *ECCV96, Proceedings of the Fourth European Conference on Computer Vision*, 347–360. New York: Springer-Verlag.



David Waltz has been vice-president of the Computer Science Research Division of NEC Research Institute in Princeton, New Jersey, and adjunct professor of computer science at Brandeis University

since 1993. He is currently president of the AAAI and is a fellow of the AAAI, a fellow of the ACM, and senior member of IEEE, and a former chair of ACM SIGART. Before moving to NEC Research, Waltz directed the data mining and text retrieval effort at Thinking Machines Corp. for 9 years, following 11 years on the faculty at the University of Illinois at Urbana-Champaign. Waltz received all his degrees from MIT. His research interests have included constraint propagation, computer vision, massively parallel systems for both relational and text databases, memory-based and case-based reasoning systems and their applications, protein structure prediction using hybrid neural net and memory-based methods, and connectionist models for natural language processing. His e-mail address is waltz@research.nj.nec.com.