

AAAI Presidential Address

(AA)AI

More than the Sum of Its Parts

Ronald J. Brachman

■ Recent thinking has it that AI, 25 years ago a unified field with a shared vision of creating intelligent machines, has devolved into a loosely connected set of distinct specialty areas with little communication or mutual interest between them. To the extent that this is true, it certainly appears to lessen the value of a centralized AI organization like AAAI and of traditional grand-scale AI conferences. But, I argue, the consequences are actually far worse: because of the very nature of intelligence, the centrifugal force on the field could thwart the very mission that drives it by leaving no place for the study of the interaction and synergy of the many coupled components that individually in isolation are not intelligent but, when working together, yield intelligent behavior. To raise awareness of the need to reintegrate AI, I contemplate the role of systems integration and the value and challenge of architecture. Illustrating some reason for optimism, I briefly outline some promising developments in large projects that are helping to increase the centripetal force on AI. I conclude by discussing how it is critical that the field focus its attention back on its original mission, led by a heavy dose of integrated systems thinking and grand challenges, and why after its first quarter century, AAAI is more essential than ever.

One of the privileges afforded the American Association for Artificial Intelligence (AAAI) president is the chance to stand in front of the entire membership of the organization and speak in an unfiltered way about whatever he or she has on his or her mind. This is a wonderful opportunity, yet a daunting one. Since one is keenly aware of the commitment to speak very far in advance, one can muse about the speech at many odd

moments over a long stretch of time. This allows the jotting of notes and the collection of meandering thoughts over quite a protracted period. But because of the sheer length of advance-warning time, it encourages one to be expansive and to note virtually anything one would like to opine about in a large forum. In my case, this freedom led to a great deal of random thinking and a fairly large pile of notes. But as the time drew near to speak, and I looked over what I had written, I found that there was almost no coherence to my many minor brainstorm. There were numerous specific things and a variety of independent research directions to consider, but no big picture. Then it occurred to me that this might actually be symptomatic of a fundamental problem that we are facing as a field and that AAAI is facing as an organization, and that the lack of a strongly unifying force might itself be a worthy theme for the address.

Between 2002 and 2005, I had the privilege of working at the Defense Advanced Research Projects Agency (DARPA) in a position that is uniquely important to the history of AI: I was honored to be able to serve as director of the Information Processing Technology Office (IPTO). In that role I had the opportunity to meet a very large variety of people with great ideas in all aspects of artificial intelligence and, more broadly, across all of computer science. While my ability to get into technical depth was limited by the sheer volume of conversations and visits, the breadth one sees in such a position is very hard to match in any other. The global perspective accrued through such extensive interactions with the community also afforded me the opportunity to contem-



The First AAAI Conference was Held at Stanford University.

plate the big picture and, perhaps more importantly (and consonant with the nature of the job at DARPA), to identify gaps in our national computing research agenda. It also occurred to me that that perspective was a very special asset to use in drafting this presidential address.

So, instead of addressing a technical topic in depth or picking on a single new direction—often the fodder for AAAI presidential addresses—I want to raise a broad issue and consider some larger questions regarding the nature of the field itself and the role that AAAI as an organization plays in AI. My hope is to encourage thinking about some things that I believe are very important to the future of the field as a whole and perhaps to start a dialogue about research directions and collaborations that in the end might bring us all back together and allow us to take advantage of the opportunity that AAAI affords all of us as AI practitioners.

A Wonderful Time to Mark Progress

The year 2005 was a momentous one in artificial intelligence, at least in the United States. It marked the 25th anniversary of the founding of the American Association for Artificial Intelligence (celebrated in the Winter 2005 issue of *AI Magazine*), and AAAI-05 was our 20th conference. The organization was started in 1980 in response to vibrant interest in the field,

which back then was served mainly by an International Joint Conferences on Artificial Intelligence (IJCAI) conference held only every two years. The first AAAI conference was held at Stanford University; it was very much a research conference, a scientific event that generated a lot of excitement. The conference was small and intimate, with few parallel sessions. There were excellent opportunities for us to talk to one another. AAAI-80 gave real substance to the organization, clearly getting AAAI off on the right foot, and it gave new identity and cohesiveness to the field.

This year—2006—has also been a big year, celebrating the 50th anniversary of the original meeting at Dartmouth College, where the name “artificial intelligence” first came into common use. Numerous events around the world, including a celebratory symposium at Dartmouth and an AAAI Fellows Symposium associated with AAAI-05, have marked this important milestone in the history of the field.

Progress since our first AAAI conference has been substantial. While each year’s results may have seemed incremental, when we look back over the entire period we see some truly amazing things. For example, a computer program finally beat the world’s chess champion. In hindsight this may no longer look so exciting (purists will say that it was not an “AI” system that beat Garry Kasparov but rather a highly engineered special-purpose machine largely made possible by Moore’s Law), but it is worth contemplating from the point of view of 1980—or, even more dramatically, from that of 1956. Looking forward from back then, no matter how Deep Blue actually worked, playing chess well was clearly an AI problem—in fact, a classical one—and our success was historic. More recently, a robotic vehicle from Stanford University, using machine learning technology, conquered the 2005 DARPA Grand Challenge, successfully managing a course of more than 140 miles over difficult terrain in less than 10 hours without any human intervention. By any measure this was an incredible feat.

Another notable aspect of AI life over the last quarter-century was the broad rise of excitement and financial investment in the field. That ranged from a wave of startups that lasted into the 1980s to significant diffusion of our technology and practitioners throughout worldwide industry. The influence of the AI community in current large and critical commercial enterprises—through, for example, text processing, speech and language processing, robotics, machine learning, data mining, knowledge management, and a host of appli-

cations of the sort that the Innovative Applications of Artificial Intelligence (IAAI) conference has highlighted for years—while accomplished perhaps with little fanfare, has been undeniable. Important systems with significant AI contributions are deployed and operating daily in virtually every industry. Small robots have saved lives, both on the battlefield and in difficult search and rescue situations. AI systems are flying in space. And through Internet search engines, every day AI directly touches millions and millions of people.

Perhaps the zenith of our field's popularity was in 1985, when we held our joint conference with IJCAI in Los Angeles on the University of California, Los Angeles (UCLA) campus. That conference had more than 5,000 attendees. The excitement in the community was palpable, and the interest from the outside, both in the commercial sector and in the press, was extraordinary. There was a great deal of national and international attention. For example, Woody Bledsoe (then president of AAAI) and I were asked to appear on a national radio talk show, where we debated various aspects of AI. Hector Levesque gave the Computers and Thought lecture in a crowded Pauley Pavilion, the home of UCLA basketball—quite an exciting experience. We probably haven't seen anything like that since, but it was a wonderful time.

Since then, we've had an "AI Winter," with a dramatic drop in funding for the field, followed by many years of reasonable growth and significant thaw. One small indicator of the current improved situation, based on our work at DARPA: starting with the fiscal year 2006 budget, there has appeared a line item (a "Program Element") in the U.S. federal budget that explicitly calls out "Cognitive Computing Systems"—very much an AI agenda. This expressed the government's intent to fund a significant budget item directly focused on artificial intelligence research, as well as at least a modest suggestion of longevity for the item. Since then, several hundred million dollars of funding were approved and spent for this area. And this represents money coming from one funding agency—it does not include current and prospective funding from places like the National Science Foundation (NSF) and National Institutes of Health (NIH) and the military services' research organizations. AI has reached the level of explicit high-level line items in the U.S. budget.

Finally, there is the undeniable general infiltration of AI-related ideas into the public consciousness. Prior to 1980 AI was an interesting topic for science-fiction writers and media pro-

ducers, with the HAL 9000 Computer in *2001: A Space Odyssey*, *Star Trek*, and the early *Star Wars* movies, which together brought robotics and the idea of intelligent androids to Hollywood. But since then, we've seen substantial growth in the public vision of interesting possible futures for artificial intelligence, including the recent film *I, Robot*, based loosely on the Isaac Asimov stories, and Steven Spielberg's direction of a provocative film that was expressly called *Artificial Intelligence*. Who among us, 25 years ago, would ever have imagined a Hollywood blockbuster, created by one of the great directors of our time, with the very title, *Artificial Intelligence*? Quite remarkable. While it is unclear how much the specific story had to do with what we do every day in our research labs, nevertheless, there it was, front and center in international popular culture: what might the far future of AI be?

On the other end of the spectrum—and perhaps more importantly—we have the well-loved Roomba robot vacuum cleaner, from iRobot. Through a remarkably modest invention, in a way that we may never have foreseen, AI has finally begun to enter the homes of tens of thousands of regular people who have no idea what artificial intelligence is about, but who know they benefit from what it brings to their lives.

In virtually every respect, our field has come a long way in the last 25 years.

Centrifugal Intellectual Force

One of the natural consequences of the growth of interest in AI in the last 25 years has been an explosive rise in the number of venues in which AI researchers can present their work. We have seen a large number of conferences develop, as well as a large number of new publications, in all of the specialized areas that at one point were all considered simply core AI. These include, among others, machine learning, knowledge representation, automated reasoning, agents, robotics, vision, planning, uncertainty in artificial intelligence, computational linguistics, and data mining and knowledge discovery. In fact, as mentioned at AAAI-05, there may be as many as 30 "sister conferences" that we can draw from for our sister conference report sessions. For quite a long time in the earlier days of this expansion, interest continued in the core conference—in AAAI itself—with people also spending time at their own specialized area conferences. But over time this seems to have changed. With strong specialized program committees to review submissions, and a greater concentration of people

... the specialization and the kind of centrifugal force pulling people away from AAAI as an organization and a conference might actually be having a much more challenging effect on our ability to do the very work of our field.

in the specialized disciplines attending these conferences, the more narrowly defined meetings have become more attractive. If funding becomes tight or if people are doing too much traveling, AAAI and perhaps IJCAI conferences are left off of the itinerary; by attending the Knowledge Representation (KR) conference or the International Conference on Machine Learning (ICML), or the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), many discipline-focused researchers are still getting to see most of the colleagues that they would like to talk to and hear about the latest technical developments in their areas.

As a consequence of this, what we hear often in AAAI organizational circles are concerns about the future of our own organization: What is its role? Why are people not coming to the AAAI conference? Should we bother to have the conference at all? In fact, if you were to look back over the last several years' worth of statements written by candidates nominated for AAAI Executive Council positions, you would probably find a majority who have focused on the issue of specialization in the field detracting from AAAI as an organization and possibly decreasing clarity of the meaning of our central conference. This is a common theme that has been discussed often in AAAI leadership circles over the last few years. There are good reasons for this evolution, and I think it is a natural phenomenon; it can be seen in other fields and in some ways it may not be something to be worried about.

But my concern here—and the main theme of my talk—is around the fact that it may be the case that in our field—in contrast with many other fields—the *specialization and the kind of centrifugal force pulling people away from AAAI as an organization and a conference might actually be having a much more challenging effect on our ability to do the very work of our field.* While a natural consequence of the maturity of the field is the “loss” of some of our subfields, this specialization and spinning off of areas of research presents us with a deep and fundamental problem.

We have been seeing great progress along specialized lines in each of the areas of AI, and some have made extraordinary gains over the last 10 years (machine learning being one of the most obvious cases). But in general it appears that while we are getting greater technical depth and greater achievements in our specialized fields, we do not seem to be getting any closer to what people might consider “true AI”—the goal that caused us to start AAAI in 1980 and that gave birth to the field in the 1950s and that, to be honest, got most of us very excited about getting into the field when we were even back in high school—the grand vision of building a truly intelligent human-level AI system. As a whole, the field just doesn't seem to be making a lot of progress in that direction, even while we make tremendous progress in our specialized areas. In my view, this is a consequence of too much focus on AI-related “components” and not on artificial intelligence itself, which is not simply the sum of its component parts.

Another way to look at this (to pick somewhat unfairly on one subfield) is to consider the fact that, no matter how outstanding our learning algorithms get, it's hard to call any of them “intelligent.” The world's greatest support vector machine is not an intelligent system that has the robustness, breadth, and common sense that even a three-year old child has—nor will it ever be. In a way, the peculiarity of our current situation with respect to AI as a whole is exemplified by a simple phrase that is commonly used in the learning area: we seem to have made a lot of progress in “knowledge-free learning.” Think about that: imagine talking to an average person on the street about AI, trying to explain that we want to build intelligent humanlike robots, and that we are making tremendous progress in “knowledge-free” learning. I suspect that that would be at least puzzling for a lay person, if not downright bizarre. Why would you handicap your system by expressly discarding anything that it had learned in the past or knowledge it had gained our could infer through other means? How could you even have the concept of knowledge-free learning in a fully integrated AI system? Yet this is common terminology, exemplarity of a common point of view.

Similarly, as much as I am personally fond of work on the semantic web, not to mention that much of my own technical work was involved in description logics that provide a foundation for web ontology languages, no matter how far we go down that road—no matter how wonderful our description logics become—they won't by themselves lead us to artificial intelli-

gence. As with isolated learning algorithms, they may have fantastic applications, they may have great intellectual depth, and they may even make some people a lot of money, but I do not believe that such specialized tracks will lead us to AI. The same can be said for virtually all of our subdisciplines, for example, machine perception and natural language processing. They may produce wonderful artifacts and great scientific breakthroughs, but they don't seem individually to be leading us to an artificial intelligence. So my concern goes way beyond the fact that specialized conferences are drawing attendees away from AAAI conferences. The centrifugal acceleration that we have been seeing, with more and more people spending more and more of their concentrated time in the subareas and at their own conferences and in their separate journals, may actually be sabotaging the future of our own field at its very heart.

Intelligence Is a Many-Splendored Thing

It has been hypothesized that whatever intelligence is (and we obviously have not been able to fully define it so far), it is a multidimensional thing. Natural human intelligence appears to be multifaceted. "Human-level intelligence" does not lie at the end of the road that focuses only on machine learning algorithms, nor will it be the result of the world's greatest natural language processing machinery taken in isolation. It is somehow a combination of these things and more. Let's think about this commonsensically for a moment: imagine that you could build a system that was fantastic at all the kinds of reasoning that people do, but without the ability to learn. Such a system, even with all of the knowledge in the world put into it right now, would no doubt appear smart in the first instant, but without the ability to adapt as the world changes and knowledge needs to evolve, it will simply become more and more stupid over time. That is, systems that are knowledgeable at one moment, but have no adaptive capability, ultimately become stupid because they just stay set in their ways and hold onto beliefs that are increasingly out of touch with reality. Similarly, systems that "learn" without knowing anything—the knowledge-free learning idea—will ultimately become very good test takers, being able to spit back just what their teachers have told them, but they will not be able to apply what they have learned to truly novel situations, which in fact is the essence of the versatility that humans exhibit every moment of their lives.

Even an average human can learn something in one context and then later use it in a new one; as a result, the average person is able to cope with many things that he or she never specifically anticipated and is not directly prepared for. But you can't do that if you just have a learner that doesn't know how to store, represent, and reason with the knowledge that it learns. In a similar fashion, perceptual systems that may appear to be excellent at dealing with pure data-driven, bottom-up data collection tasks, but that are not guided by cognition, will not work very well in many complex situations. When the input is multidimensional and full of interesting tidbits, they will get distracted by things that are not important or are not material to the task at hand. They may not recognize that something is an anomaly, or on the flipside, they may end up paying attention to something that is a completely unimportant anomaly. When you look at human and animal systems, perception is guided very strongly by mechanisms in the higher parts of the brain in some magical way that we currently don't understand, but that guidance is very critical. Despite the fact that our sensors are independently quite competent, they are not just data-driven transducers doing blind pattern recognition—they are much more effective when working together and, in fact, when guided by common underlying cognitive mechanisms that can synthesize their various outputs and use expectations to cut through noisy, distracting data.

Note that success or failure is not a matter of the quality of the individual components. We are able to use flawed parts—our vision and hearing systems are less than ideal, our natural language generation systems are imperfect, our judgment is flawed—and yet, nevertheless, as a whole, as something that in a way is more than the sum of its parts, we do very well getting around in an unpredictable world that throws things at us that we can never anticipate. So, even the ultimate end products of each of our specialized disciplines done perfectly well are not likely to yield an intelligent system without us doing something fundamentally different. This is perhaps the key point of this treatise: AI work done in separate subdisciplines—even that resulting in practically flawless components—is not itself "AI." We must consider the integration and synergies of components in an overall system to really approach some form of artificial intelligence.

Now the idea that intelligence is actually built out of many parts and is not a single, monolithic thing is not a new idea. Many people have talked about this over the years. One



of the things that we did during my recent tenure at DARPA was to look at ways to test intelligent systems that have multifaceted capabilities. It occurred to us—and I want to thank both Dave Gunning and Paul Cohen for this idea—that we may want to test intelligence skills and capabilities in machines like we do very broad athletic abilities in humans. Someone might be the world’s greatest shot-putter, but world-leading performance in the shot put alone doesn’t seem to be indicative of whether or not the person is, in fact, a fantastic *athlete*. They might be the physical equivalent of an *idiot savant* and have only a single skill that doesn’t translate well into any other form of athletic performance. To determine who is the best overall athlete—the most “athletic”—we typically use the *decathlon*. We run a set of events that themselves aren’t connected in any obvious way (throwing the javelin may have nothing to do on the surface with running hurdles), but nevertheless, there are muscles and coordination and general athletic conditioning that somehow support all of these. Therefore, by testing a handful of different skills, we come to the belief that someone who may not even be best in any individual event, but somehow is best in the aggregate, is truly the world’s greatest athlete, exemplifying what we consider to be athleticism. It may be that at the moment, given our crude understanding of natural intelligence, this is perhaps the best (or maybe the only) way to test intelligence. Being “intelligent” may be the same kind of amorphous mix that being “athletic” is, or alternatively, given our poor understanding of it, a multifaceted test may be the only way to approach the measurement of intelligence.

Flying in Formation

Many others, largely from the psychology community, have talked about intelligence as a multifaceted thing, and there are “multiple intelligence” theories. Similarly, some in AI (for example, Marvin Minsky in *Society of Mind* (New York: Simon & Schuster, 1985) and subsequent writings) have talked about intelligence being multifaceted. Let us grant for a moment that indeed intelligence is not a single, monolithic capability. There is one more critical point: intelligence is not created by just mixing together the individual facets. Indeed, intelligence takes many different individual, special capabilities—like reading and thinking and learning and seeing and hearing and talking—but, even if we have all of those pieces each done extraordinarily well, if we just toss them together we are not likely to get an intelligent system. My sense is that intelligent systems are not simply the mere compositional sum of their parts. Individual senses and cognitive abilities don’t exist separate from one another and don’t simply drop their independent conclusions into a centralized data store that they can all read from independently (the appeal of blackboard architectures notwithstanding). If you start to examine what people actually do, you see very deep and complex connections between these different capabilities.

Unfortunately, these relationships are the things that are often missed in our specialized conferences because in order to analyze and understand them, specialists in one area would need to reach out to communities that have their own specialized conferences that are intellectually quite far away. For example, natural language understanding in humans, which is often addressed in computational linguistics conferences, really does seem to involve *understanding*—as tautologous as that sounds, the focus of most natural language processing work these days is *not* on understanding, but on doing as much as one can *without* positing an understanding component. Understanding involves world knowledge, reasoning, and the ability to do inference and see the consequences of things that you learn and apply them to new situations, which soon gets quite out of the realm of what you typically see at a computational linguistics conference. As I have said several times now, learning produces knowledge as an output and that knowledge needs to be used in reasoning components; without this, reasoning systems can chug along and do work with what they’ve got, but they end up being pretty stupid unless they can take advantage of newly generated knowledge from learning mechanisms. Additionally, real-world

learning in the kind of normal everyday situations in which we find ourselves is heavily dependent on what we already know, or expressly seek to learn because of other things we are thinking about. Further, as I have also said, perception doesn't really operate independent of cognitive state.

As we examine more and more of the pieces of an intelligent system—the separate events in the cognitive decathlon if you will—we find that we still need to account for the magical way in which they all have to fit together in order to create something that is more than the sum of its parts. Perhaps the best way to get this point across is to quote someone I used to work with (Dado Vrsalovic), who, in a different but perhaps comparable context, was fond of saying the following: “100,000 spare parts flying in close formation is not the same as an airplane.” I am not sure how you launch 100,000 spare parts to fly in close formation, but imagine taking all the parts in a Boeing 777 and putting them really, really close to one another, but not totally integrated in the way we like to see our airplanes integrated when we fly on them. Now when you launch this thing, it isn't going to stay up in the air very well and, in fact, it isn't really an airplane. There is no *architecture*—an integration that needs to be taken into account to take the parts and turn the whole into something that really serves a broader function.

So somehow, even in the decathlon, while we test skills separately, they are all being done by the same person with much overlap in the subsystems that are used in the various events. Running hurdles and doing the long jump clearly use related leg muscles; throwing the javelin and pole-vaulting both use a running start and significant arm action. Imagine though, a field that had emerging specialized journals and conferences that focused totally on the mechanics of javelin-throwing. Members of the Hurdle Society and attendees at the International Pole-Vaulting Conference would no doubt be in the dark about what came to be discussed at the javelin meetings—and in fact might declare those fields to be so far from their interests that they were no longer interested in attending general athletics conferences where there would be boring and irrelevant papers on javelin-throwing, sprinting, and high-jumping. If the experts in those specialized areas were interested only in creating machines that could star in their individual events, they might be able to get away with this separatism. Even then, of course, the javelin folks might miss an important development in the discus-throwing literature that

could provide critical insight and might even help their subfield take a quantum leap ahead. But much more importantly, if the genesis of the separate subareas was a once-exciting field that had originally aspired to create a general artificial *athlete*, able to compete competently in *many* different events—including novel athletic competitions that had not previously been on the program—then the subareas going their own separate ways would fly directly in the face of the original overarching goal. It seems to me that the way to build a competitive artificial decathlete is not to lose the focus on the athlete as a whole by working independently on the subsystems needed for specialized events. Certainly the leg and arm systems would need to be studied in detail and emulated in machinery; but their interactions to produce the rhythm of a long-distance runner or the coordination necessary to achieve a successful pole vault would need to be a mainstream concern of the field as a whole. Without attempting to study and build a whole body with integrated musculature, where would balance, strength, endurance, and coordination come from?

Now this is admittedly a naïve view, and I am no physiologist, but I would suspect that if you looked closely at how athletic skills are implemented in natural systems, you would really need to pay attention to the cross-coupling of different muscle and skeletal components, not to mention the overall cardiovascular infrastructure—and the key point here is that it is the same thing with intelligent systems. “Learning” isn't really *learning* in intelligent systems until it is tightly connected with reasoning. Perception doesn't mean much unless it is coordinated with cognition. Language essentially involves thought and understanding.

The implications for AI should be obvious. If we simply keep our separate components flying in close formation, we won't ever really get an AI airplane. And if we build machines that win races in specialized areas, we won't get a decathlon winner or an athlete capable of succeeding at a newly invented event. We need to think about how all the parts of an artificial intelligence should work together, and how they need to be connected, or we simply won't get one. So to my mind, what we really need to do as a field is to spend a lot more time focusing on comprehensive, totally integrated systems. It doesn't look like we are going to get there from here if we keep going in the direction we are going, especially given the centrifugal force that is drawing people away from one another into their separate subfields.



Mutual Influences

Now, why hasn't there been a lot of work on integrated systems? In fact there have been some wonderful role model projects, but by and large, if you look at AI, most of it is explored in specialized venues or in narrow, specialized sessions at conferences. As a result we don't spend a lot of time talking about pulling the pieces together and creating the magic glue that turns them into intelligent systems rather than just sets of pieces. First and foremost, it is clear that building full, integrated AI systems is extremely difficult. It takes a lot of talent with a lot of breadth. It requires sizable teams, software engineering skills, and even project management skills—the kinds of skills that are not deciding factors in hiring university professors. Graduate students are really not taught this, and it is often the case that many engineering and architecture considerations are deemed “below” some scientists. It is notoriously difficult to get AI systems papers published in key places.

Nevertheless, these things are absolutely critical to our ultimate success. As a field we don't necessarily have a lot of these skills around. Building large-scale integrated systems is really, really difficult. It is also very messy just because of the essence of what's involved. Everything

in the system can ultimately affect everything else. And that makes building and debugging the system very hard.

Three Focal Points

Before I get into specifics about some role model efforts, I want to talk briefly about three relatively underdiscussed issues that I believe are important. One is the notion of *architecture*. Another is what we might call *knowledge-rich learning*—somewhat the opposite of the knowledge-free learning that I mentioned earlier. The third is the nasty, unpopular topic of *metrics and evaluation*.

Architecture

We are all familiar with architectures in AI systems. We often see block-and-arrow diagrams in papers, and in fact, there is a pretty substantial amount of work largely driven from the psychological side of things around what are called “cognitive architectures.” In my view, however, there is still significant room for work in this area as applied to integrated AI systems. The cognitive architecture literature has tended to focus strongly on psychological validity, matching what goes on in the human brain. Since our concern in AI is not just the replication or understanding of humans, but in fact a broader space of possible minds (thanks to Aaron Sloman, a philosopher and long-time researcher in the field, for that description), I think we need to expand our thinking about architectures that are adequate to support integrated AI systems.

A potentially fruitful question to ask is whether we can learn from architecture work in other disciplines. If you have ever been involved with a large-scale software system effort, you'll know that architecture and systems engineering are critical parts of the equation. These arts are practiced by specialists with substantial technical education in the area, and hard-won experience is crucial to success. Architectures are not just boxes and arrows put together by the seat of the pants. System architecture as a discipline has its own principles, and we may be able to learn a lot by taking a broad look at the literature on software systems architecture, hardware architecture, and even building architecture. It would be interesting to see how well-known principles of good architecture might apply to AI.

For example, here are a few principles (courtesy of Gregg Vesonder) that you might learn about good architecture if you were taking a class in software engineering:

A good architecture is one that is amenable to

implementation of a skeletal system with minimum functionality overall in order to test the communication paths between the components. One way to get started is to build stubs for all of the parts of the system without making them rich and fully functional; this allows you to figure out whether the APIs and the communication paths are right before you invest in full implementation of the parts.

A good architecture would feature well-defined modules with functional responsibilities allocated on the principles of information-hiding and separation of concerns with well-defined interfaces.

A good architecture should feature a very small number of interaction patterns. Modules that produce data should remain separate from modules that consume data. Good architectures avoid doing the same thing in different ways and in different places.

Principles like these are familiar to computer scientists, but the kind of formal architecture reviews necessary to successful large software projects are not common in AI work. I think we have some important things to learn from the way architecture discipline is applied in more mainstream software projects.

On the other hand, we also need to consider how building an intelligent system might be different from building a more conventional software system. It certainly appears to be the case in the human brain that the architecture doesn't look like it does when we construct systems with simple components and do information-hiding and separation of tasks and only do one thing each in one place in a system. Neural architectures hypothesized by brain scientists look quite different from those of conventional software systems. One critical element here is the fact that the one role model that exists in the world that we've been looking to emulate was not a designed artifact; the brain *evolved* over millennia, and in fact, what you get with an evolutionary approach is very different and potentially much messier than something you would do with a very good engineering design. So this leaves us with an interesting conundrum: as we build large integrated AI systems, it would make sense to follow conventional architecture discipline, but we generally look for inspiration to something that evolved in an organic way and whose currently discernible architecture is very different from that of the systems we typically design.

Knowledge and Learning

Humans learn, it is clear, in many different ways. We learn by being taught, we learn by trial and error, we learn by mental simulation and explanation, we learn by reading, and we learn

in a host of other ways that may or may not all be implemented differently. When we think about learning, though, the goal is to produce knowledge or resources that we can use in dealing with novel situations later in life. So what we tried to do when I was at DARPA, and what I think is important for the field in general, was to move back to the more traditional AI view of learning. Here I'm not talking about Machine Learning with a capital "M" and "L" as a sub-field, but good old learning as done in core AI systems, where, for example, we might have systems as capable as humans are of learning from very few examples (including just a single one). Learning also is profoundly important in getting around in life when we use it in what we might call a "transfer" way—for example, absorbing things we're taught in a classroom and actually finding out many years later that they apply in unanticipated ways to real life. We learn things in one context that work out to be very useful in a different time and context.

Another phenomenon in learning is "laddered learning" or what Dan Oblinger has called "bootstrap learning," where learning works best when principles are presented in a structured approach, so you learn one "rung" at a time. We know that when things are presented to us in school in the right order, we can actually learn incredibly complex concepts, whereas if we just jumped to them right away, we could never learn them. So we know pedagogically it is very important to ladder things, that is, to master one rung before we jump to the next rung, yet that's not how we are currently approaching research in machine learning.

Consider also the difference between bottom-up, or data-driven knowledge acquisition, and top-down, or goal-driven simulation and acquisition of knowledge. These two approaches come together in learning. For example, when reading you don't just read the words; rather, you read with expectations about what will come next. You can complete a sentence even if you hide the last word in the sentence and your interpretation of the words you see is influenced by what you've just read. Your forward-looking expectations influence how you interpret the next sentence, and what topics might be presented in the next paragraph. This implies a potentially complicated architecture underlying reading that we ought to include in our learning research.

A final issue on the subject of learning systems is how to tell whether a system has learned and whether its learning can be improved.



Deep Space One Encounters an Asteroid.

Photo courtesy, NASA.

All in all, these notions of learning are quite different from the main current thrust of machine learning research. This is ironic in a way, since as far back as the first AAAI conference these concepts were posited as the basic foundations for learning.

A Word on Evaluation

Before working in the government, I, like many of us in the field, really hated the idea of metrics. It was too difficult to figure out quantitative metrics for knowledge representation and reasoning or concept learning, and when you did, they tended to distract from the main line of work. But since I started working with others on larger integrated AI projects, and securing funding for the field, I have found that evaluation is an absolutely critical ingredient of our research program. I believe metrics and evalua-

tions are requirements for the future of the field, both for our own sake and also for the sake of those who oversee our funding. In particular, in the United States it is very important to show Congress that we are actually making forward progress—even if our goals are 20, 50, or 100 years in the future. It no longer matters what our subjective description is of how much better things are than they were five years ago. Rather, it is important to understand in a measurable and repeatable way where we are heading and how we will get there.

Now this is very challenging. It's hard enough to apply meaningful evaluation designs to AI component technologies, just given the nature of research. But the difficulty is magnified when you take to heart the message that we need more work on integrated, comprehensive AI systems. How will we evaluate intelligence in a machine when we don't even know how to measure intelligence in people? But this challenge alone should prod us as a scientific community to come up with new ideas and approaches to the issue of evaluation and metrics.

Some Role Models

I am encouraged to note that there are some good examples of the integrated, systems approach to AI. I want to highlight briefly a couple of projects that I think could become role models. One involves work in the area of space exploration at the National Aeronautics and Space Administration (NASA) and at the Jet Propulsion Laboratory (JPL). The other is a DARPA project called the Personalized Assistant That Learns.

One of the most exciting premises of NASA-style space exploration projects is the fact that the environmental constraints and mission goals are such that AI must be a fundamental ingredient. Simply put, you can't teleoperate a remote vehicle in dangerous or rapidly changing circumstances when there is a 30-second (or worse) radio transmission lag. In these situations the vehicle has to care for itself in a fundamental way. Dan Clancy presented an invited talk at AAAI-04 in which he talked about the vehicles we are sending to Mars and deep space. Dan described how these machines monitor and maintain their health on a constant basis. They may get hit by meteorites, they may experience power failures, they may pick up dust in moving parts, and they may have faulty electronics—a host of things that require real-time observation and response by the vehicle itself, because there may not be time to get a signal back to earth before it's too late.

In addition to keeping track of their own health and other prosaic things, these vehicles need to assure the success of their science missions, following plans for collecting data and getting them back to earth—and changing those plans on the fly. Here we have a set of interesting constraints from a nasty set of real-world problems where AI—integrated AI—is fundamentally necessary. There are some really fantastic projects completed or imagined, ranging from the remote agent mission launched on the Deep Space One probe a few years ago to something called the “Autonomous Sciencecraft Experiment,” which has been running on a continuous basis for some time now. It is clear in this kind of application that planning, execution, perception, and action are integrated and multilayered. Further, because subsystems can override each other and opportunities for conflict abound, the architecture of the integrated system and how it carries out its intelligent activity are of absolutely fundamental concern. You can’t send a great camera, a great pattern recognition system, a great set of wheels and motor controls, and a host of other separate parts into space and expect those “spare parts” flying in formation to carry out the mission. They must be deeply integrated. One has to think in advance about all critical and dependent interactions between the parts of the system. In other words, the architecture is the overriding concern. And it is fantastic to see deployed AI systems of the sort that we all know and love actually running in real time, live on \$100,000,000 NASA space missions.

At JPL, Steve Chien and others have pointed out how these issues come into play in the design and implementation of their systems. Steve notes that having multilayered systems, coupled with the fact that there are different priorities at different times, means that we need to worry about redundancy, safety checks, real-time considerations, priorities, and constant change and adaptation of plans based on the real world’s effects on the mission. And as if that were not hard enough, systems like these can’t really be fully tested until they are launched. To run a realistic test on earth to simulate what the mission would be like in space for four months or five years is totally impractical, if even possible to design. And so we have some very interesting challenges to face when we build a true AI system that’s going to be deployed in a real-world, real-time, mission-critical environment.

The second example involves some major efforts under the Personalized Assistant that Learns Program at DARPA. One of them is SRI International’s Cognitive Agent that Learns

and Organizes (CALO)—an artificial secretary, one that aspires to do virtually all of the kinds of things that a good human secretary can do in the office. What CALO is trying to emulate is not a Ph.D-level physicist, but rather, the seemingly mundane tracking, learning, and reminding aspects of a good secretary who can adapt to real-world circumstances, improve over time, become personalized to the person he or she is supporting, and take into account the many small things that really make everyday life challenging. This is of course a very different context than putting a rover on Mars, but equally challenging and exciting from a core AI perspective. The CALO team has identified six very high-level areas of competency in which the system must learn and succeed: organizing and managing information, preparing information products, observing and managing interactions, scheduling and organizing in time, monitoring and managing tasks, and acquiring and allocating resources. The multi-institutional team has been working for several years now on a very wide variety of AI subjects, trying to put this all together into an integrated system.

To build a CALO assistant, SRI International has served as the systems integrator for a number of technology specialists. Importantly, from the very beginning SRI instituted a systems approach that required, among other things, software engineering discipline, a chief architect, and a release schedule for software that will be tested as a system and used as a prototype in DARPA’s evaluations. Of course, a great deal of fundamental research work for this project is being tackled in many wonderful universities. But CALO is more than the sum of these individual research efforts. It starts with an architecture and a design approach that brings all of these pieces together in an integrated AI system. The architecture allows specialized learning methods, created by individual scientists at different institutions, to be inserted by using common APIs and built from common core learning algorithms. These modules interact in a well thought out, well-engineered way using a knowledge base, reasoning methods, perceptual apparatus, and other things that go into a complete integrated AI system. It is the architecture for bringing these things together where they can learn from one another, interoperate, and deal with inconsistencies that is the essence of this project, and one of its greatest challenges.

The results so far are very promising. After the first year of the project, the SRI team developed a flexible, true AI architecture with many components that could interact in at least a



Illustration courtesy, SRI International.

CALO: Cognitive Assistant that Learns and Organizes.

minimal way to accomplish a very wide variety of things—ranging from moderating meeting activity and taking notes, to preparing briefing packets that you can take in advance to a meeting. There are also obvious things such as helping with putting mail in folders and dealing with calendars, and so on—the kinds of things you would expect from an office assistant. But the exciting thing to me about this project is the deep integration between the many components and the different types of capabilities that CALO brings to the table. Tom Dietterich recently noted some specific ways in which the CALO model has driven new types of research, outlining to me some really interesting novel challenges to people doing work in machine learning. In fact, one of the more exciting things about this project is the way that research in machine learning has been affected by the requirement of building this integrated system. One of Dietterich's observations is the fact that learning and reasoning must be combined in a tight fashion in CALO. This means that, especially when you've got multiple learning systems, when a prediction is made and the world either verifies or refutes that prediction, or the user provides additional feedback ("No, that was the wrong document to give me for that meeting"), the system must keep track of the new knowledge and assign credit to the component or set of components that collaborated to make that recommendation. But there may not be an obvious path back, and in fact, different components can collaborate to come up with a recommenda-

tion, making it very difficult to assign credit. So here is an important, somewhat novel problem in machine learning that might not have surfaced if the focus were simply on a boosting algorithm or support vector machine. Similarly, this kind of integrated, long-lived system needs to formulate its own learning problems. Because it is running autonomously and is trying to be helpful, and because it knows it needs to improve, it needs a plan for acquiring, in an unobtrusive fashion, a set of labeled training examples and to evoke feedback from the user without being too annoying. These are things that people just do naturally; CALO has to figure out how to do this using deep interaction between a planning component, perception and action, learning, and reasoning technology that are all part of a complex, always-running, always-evolving system.

Not too long ago, Dietterich observed to me that in working on CALO, he was now talking with people he hasn't spoken with in 25 years. In 1980 at the first AAAI conference, we were all working on the same thing—it was AI. While we were individually focused on different pieces, we were at the same time working together and communicating. Twenty years ago, learning people didn't totally divorce themselves from knowledge representation people, and natural language people didn't go off and spend their time at their own conferences. Today, projects like CALO and the NASA-related projects have rekindled the kind of collaboration and communication that we had at the beginning. I hope this is a harbinger of the future—a way to bring the field back together and help us develop true artificial intelligence systems.

Versatility, Purposeful Perception, and Doing Something Reasonable

One of the historic frustrations in our field is that when we have built something that we believe is a successful example of AI, as it gets press and gets analyzed, it seems no longer to be AI. Instead, it becomes some form of engineering or just a tour de force. I believe that this is not really a great mystery, but that we usually bring it on ourselves by solving specialized problems with *components* of intelligence, not complete, integrated intelligent systems. Deep Blue is a classic example: what succeeded in the very specialized domain of chess wasn't an AI system embodying general intelligence with perception, knowledge, reasoning and learning, but a set of very special purpose algorithms that allowed the system to succeed at a

specialized task. We see this all the time with the specialized tasks that AI addresses. (It's fair to say that chess is such a specialized problem that in the end we discovered it didn't really need general intelligence. But if so, we certainly need to find some challenge problems for the field that *do* need true intelligence.) So we need to ask the question: "what really is AI and what is intelligence about?"

To my mind the key word may be *versatility*. (I'd like to thank Rich Doyle for first getting me to think about versatile systems when I was at DARPA.) Humans are versatile: they can take on new missions all the time; they can learn to do new things. Nils Nilsson, in a paper talking about human-level AI and grand challenges for the field, talks about educable, or what he calls "habile," systems. Versatile or habile systems embody the real value that true intelligence brings to the table—that is, the ability to do almost anything, maybe not brilliantly or wonderfully, but at least adequately well, and to be trainable to do new, previously unanticipated things. I think we need to pay more attention to this notion of versatility in our field. When the Rover completes its planned mission on Mars and yet is still functioning, it would be wonderful if it could adapt to a new mission—be "re-purposed," if you will—merely by us explaining to it what was needed or educating it didactically on a different class of missions. A human could easily do that. However, this is almost impossible to do with a classically structured NASA spacecraft because the mission is built into the architecture of the system.

General intelligence, or real AI, would allow us to describe, in a declarative way, a brand new mission and a recommended approach to that mission that the system would understand and could then take on. In fact, one of the earliest thoughts on a true artificially intelligent system was McCarthy's proposal of an "advice taker," back in the late 1950s.¹ While we've never really built such a system, McCarthy's original idea remains much more about true AI than most of the "specialized AI" systems we have built since then.

Another area that I believe is well worthy of more attention is what Manuela Veloso and others have called "purposeful perception." This is not passive absorption and purely bottom-up processing of sensor inputs, but rather perception guided by and influenced by higher-level cognitive processes and current needs. What in the perceptual field needs paying attention to, and how sensor inputs should be interpreted, is clearly driven at least partially in a top-down way in natural systems. This kind of architecture is worth understanding and

emulating. In the CALO project, for example, the perception subsystem should be influenced in an appropriate way by the overall current priorities, upcoming tasks, things that have been learned, and similar cognitive aspects of being intelligent.

Finally, a great mystery of naturally intelligent systems is that when presented with new situations, we are almost always capable of doing something *reasonable*. Our responses may not be ideal, but we don't simply go catatonic when we are faced with a new situation. We, generally speaking, muddle our way through and most often arrive at an adequate outcome. Afterwards we can reflect on what we've done and learn from our experience and do things better the second time. This is a concept we can all understand, at least informally, but reasonableness hasn't been thought about a lot in AI. Why is this? As with the other issues just discussed, I believe this escapes serious study because it is another virtue of an integrated, comprehensive intelligent system, not just a collection of spare parts hanging around together. And we're still paying almost total attention to the spare parts. Versatility, purposeful perception, and reasonableness are all exemplary characteristics of real, honest-to-goodness intelligent systems, and all seem to require the kind of integration that I have been talking about.

Measuring Progress

Once we build an intelligent system, how do we know if it's any good? How do we tell if it is doing well, especially if its behavior is kind of reasonable, it is not great at any one thing, and every piece of the system influences every other piece? How do we actually measure progress? Again, the CALO experience points in a promising direction. For the project, we hired a full-fledged, independent evaluation team to identify the task challenges for the system and then decide on the right measurements to be applied year after year—thus giving essentially the same test each year, but without the system being able to cheat because it saw the test the year before. In a way, the project has developed an SAT-like test for the system by creating a set of parameterized questions, or question templates, that can be varied from year to year, but that require precisely the same capabilities and intelligence to answer each year. For CALO, we can ask about a new set of meetings or tasks for the office assistant each year but use the same testing to end up with a quantitative score of performance and track progress from one year to the next. What really matters is whether we can

show that a system can learn, that it will ultimately do much better than an engineered system because its knowledge growth and its improvement in skills just outstrips what the design team can do as programmers.

AI Grand Challenges

If our goal is to build a machine that has some form of true, versatile intelligence, and yet the field is not progressing much in the right direction, how can we stimulate more of the right kind of work? Certainly having a system “situated” and interacting with the real world is an important element, partially because whatever AI system we build will need to work in a real context. But there is more. The real world is messy and unpredictable, so I think we need to spend more time building systems that we experiment with in the wild, not just in the lab—we need our AI systems out there dealing with the messiness of the world, dealing with noise and misdirection and things that we wish we generally didn’t have to deal with.

Another important element of a refocusing of the field is to move AI to be thought of more as a system science—not a handful of disconnected subdisciplines, but an overarching view of how those disciplines and the resulting technologies they create interrelate. In fact, I think we need to find room in our conferences for all aspects of computer science that can contribute to and benefit from the needs of complex AI systems. How can we do this, given our current commitments to our universities and companies, our students, and our careers, which tend to demand simple allegiance to fields defined by existing conferences and journals? What kind of game-changing event might create the motivation we need to leapfrog to a new paradigm? I suggest that we launch a new series of “grand challenges.” Such challenges, like President Kennedy’s bold promise in 1961 to put a man on the moon (and return him safely) by 1969, can grab the attention of a broad group of people. They present hard problems to be solved—hard enough to require lots of work by lots of people, and when defined right, they move fields in significant leaps in the right direction. The really good ones over the years are interesting because they have really demanded the kind of integration that I have been advocating. In my view, with appropriate sets of challenges, we might actually generate a tremendous amount of new excitement in the field of AI systems, which doesn’t really have its own separate sub-area or conference.

Criteria for Successful Grand Challenges

There are aspects of creating such challenges that require some serious thought. First and foremost, we need a problem that demands a clear and compelling demonstration of AI. Such problems, as I’ve intimated above, will need to stress versatility, adaptation, learning, reasoning, and the interaction of many aspects of intelligence. Next, a successful challenge cannot be gameable in a way that somebody could win the prize with very specialized engineering that doesn’t get us closer to an AI system. This requires extensive thinking about the ground rules for the challenge. Another thing that is really critical to a successful grand challenge is that the measurements—the success criteria—are well-defined, easy to articulate, and understandable to any lay person who might participate. For example, the success criterion for the DARPA Grand Challenge autonomous vehicle competition in the desert was simple: travel on the ground along a prespecified route from point A to point B in less than 10 hours, with no human intervention—period. You don’t have to say much more—people know what you want and what it takes to succeed.

Another key aspect of a good grand challenge is to make the problem hard enough such that it is not clear that it can be solved and to allow partial successes to lift the general state of the art each year. Here, failures along the way can point the direction we should go to improve things—a critical aspect of good grand challenges. (Thanks to Paul Cohen for the notion of failure being “diagnostic” as an aspect of a good grand challenge.)

Of course, the challenge has to be ambitious and visionary, like putting a man on the moon or cracking the human genome or beating the world’s chess champion, but it can’t be unrealistic. As much as I would like to put out the challenge of building a Star Trek–like transporter, I don’t believe such a thing would be possible in my lifetime (if ever). People will get frustrated if they don’t see their way clear to possibly solving the problem and winning the prize. We can debate what the right time horizon is, but we certainly want to have solid faith that someone will succeed in somewhere between, say, 5 and 15 years, with reasonable doubt about solving it in the first 5 years. That’s what makes it really fun.

A successful AI grand challenge would need to be compelling to the general public, generating good press and building general excitement. The public has to find it compelling and exciting enough to follow the stories in the press and be rooting for the AI researchers to

succeed. Finally, a successful challenge has to be strongly motivating for the research community, one that gets specialists talking with each other and exploring new collaborative paths to success, asking new questions and questioning past beliefs.

The aforementioned criteria should be applied rigorously to grand challenge suggestions. Hopefully we will end up with not just a laundry list of possibilities but rather a compelling, exciting quest that will attract researchers and lay people to find a solution. I believe AAAI could play a very important part here.

Some Examples

What specific problems might qualify as high-quality AI Grand Challenges? Back when I was at DARPA we held a workshop to identify, debate, and refine some ideas. Here, briefly, are some of the ideas that were still on the board at the end of the workshop.

Truly read and understand a book. Here, the challenge would be to build a system that can read a book (a textbook, say) and successfully answer the questions at the end of a chapter or, better yet, sequential chapters. To make such a challenge compelling, it probably has to be the equivalent of what a student does in a high school science course, with some details and ground rules appropriate to computer programs. But pretty much, that's it. We would have to think through the ground rules and very carefully plan a nongameable way to administer the test. But the overall problem is intuitive and challenging. We could easily imagine a poster announcing the challenge.

Take the Scholastic Aptitude Test (SAT). Here each system would take the SAT on the same Saturday morning that human students do in virtually the same context and would be required to end up in the 90th percentile of high school student performance. Again in this case, it is very simple to describe the challenge and it would be very clear to even the general public how monumental an achievement this would be. Many people complain that the SAT is not a "good" or adequate test for humans of high school age. But no matter how flawed we all think the SAT is, we use it, our kids take it, and it matters for getting into college. It is fairly broad and involves a substantial amount of background knowledge and the ability to apply learned skills in novel settings. Wouldn't it be exciting if a computer program could, under the same ground rules as the students, actually do well?

Write an essay that is graded well by a high-school teacher. Picture your typical school



research paper assignment: write a short essay on an assigned topic, first doing some research using defined information resources, and then producing a coherent paper in a finite period of time. For example, I may not know much about the French revolution, but if I were given two hours and access to the world wide web, I could probably put together at least a plausible two-page essay on the causes of the French Revolution, or what happened at the Bastille, or the origins of "Let them eat cake!" Why shouldn't a machine be able to do this? Why don't we have computers try to write essays just the way we have kids do it, have the essays graded by human teachers in exactly the same way, and then as soon as one of them gets an "A," celebrate a monumental step in the history of AI?

Be a “search engine” in the real world. Why don’t we build search engines for the physical world? In fact, why don’t we create a grand challenge treasure-hunt adventure? Imagine we are going to take mobile robots entered into the competition to a previously undisclosed physical location. It could be a house, but one that the entrants have never seen before. We tell them to find a \$1,000,000 treasure in a chest in one of the rooms (or some other challenging physical and mental quest). Make the circumstances difficult enough such that it would test a human. In the end, the treasure would be much more than watching the robot bring out the chest—the real treasure would be the advances in AI systems and science that are brought to the world.

Of course, the sky is the limit here. You can imagine a very wide variety of challenges, which would be successively harder and in some cases would ultimately appear to take truly an integrated AI system to actually solve a problem like this and win the grand prize. Forget a vehicle that can drive itself across the desert to Las Vegas—let’s challenge our colleagues to build intelligent machines to drive to Las Vegas from anywhere in North America, find a reasonable hotel for a night, check in, and then win some of its designers’ money back in a casino, walking to the cashier with its chips, and depositing money in the right bank account....

Conclusion: The AAI Challenge

This somewhat meandering discussion leads me to a very simple and obvious conclusion. Who will lead the way to these new, integrated systems? Who will bring together our apparently divergent research specialties? And who will lead the charge in creating, motivating, and running a set of AI Grand Challenges? Clearly, if we are going to focus more on integrated systems work, what you might call true AI, we still need a venue for exploring and sharing our results, talking to one another, and also discussing and running these grand challenges. I don’t particularly think that this should be a government activity, although if we can convince DARPA to sponsor another grand challenge that demands more integrated intelligence, that would be great. I think these things should be the province of the researchers and scientists and leaders in the field.

The obvious place to do this is, of course, AAI. In some ways, this is exactly what the organization was founded to do more than 25 years ago: to be the central place where all of us come together, in a conference setting, in

print, and as an organization. It should sponsor activities to bring together people who ultimately need to collaborate, and create an atmosphere where those researchers could work together closely to build truly integrated AI systems. This movement back to the roots of AAI has to start with the AAI leadership, and with each of you.² We all need to do what Tom Dietterich observed that he was forced into doing: all of us should in fact start talking much more to colleagues we haven’t spoken to in 5, 10, or 25 years. Only through such personal, substantive collaboration, facilitated through AAI’s operations and meetings, will we be able to reach our goal of building the kind of AI system that was the heart and soul of the field when AAI was founded in 1980, and was the dream of the founders of the field a half-century ago.

Note

1. See John McCarthy’s “Programs with Common Sense,” Teddington Conference on the Mechanization of Thought Processes, December, 1950 (www-formal.stanford.edu/jmc/mcc59/mcc59.html).
2. I am pleased to note that *AI Magazine* (27[2]:2006) recently featured a special issue with the theme, “Achieving Human-Level Intelligence through Integrated Systems and Research.”



Ron Brachman is the vice president of worldwide research operations at Yahoo! Research, the advanced research arm of the worldwide leader in Internet services. He earned his B.S.E.E. degree from Princeton University (1971), and his S.M. (1972) and Ph.D. (1977) degrees from Harvard University. Brachman has made numerous important contributions to the area of knowledge representation, including *Knowledge Representation and Reasoning*, a recent textbook written with Hector Levesque. Brachman started his career at Bolt Beranek & Newman, spent several years at Fairchild/Schlumberger’s Lab for AI Research, and, having developed a world-class AI group at AT&T Bell Laboratories, moved into senior research management jobs at Bell Labs and AT&T Labs. He served as the director of DARPA’s Information Processing Technology Office from 2002 to 2005, and there developed IPTO’s Cognitive Systems initiative, which brought hundreds of millions of dollars to the national research community. Brachman was program chair for AAI-84, and was the president of AAI from 2003–2005. He is a founding AAI Fellow and a Fellow of the Association for Computing Machinery (ACM). At the International Joint Conference on Artificial Intelligence in January of 2007 he will be awarded the Donald E. Walker Distinguished Service Award.