Crossroads
The ACM Student Magazine

An Interview with: Drew McDermott

by ***Kentaro Toyama***

Drew McDermott graduated with a combined B.S./M.S. Electrical Engineering degree in 1973 (``I had to pretend to be an electrical engineer for 4 years as an undergrad''), and wasted no time earning a Ph.D. in 1976, in the newly-formed Electrical Engineering/Computer Science department, all from MIT. He then immediately joined the Yale faculty and rapidly ascended to the top of the ivory tower.

He's well-known in AI circles not only for his extensive work in logic, planning, and robotics, but also for his blunt public appraisals of the state of AI research. In 1976, he wrote an essay titled ``Artificial Intelligence Meets Natural Stupidity,'' in which he chastised AI researchers (himself included) for various forms of wishful thinking that allowed one to interpret a (sometimes unimplemented) program's relatively meaningless data manipulation as meaningful, intelligent thought. In 1987, he effectively renounced an area of research to which he had devoted many years of work: ``A Critique of Pure Reason'' expresses his disillusionment with attempts to base common-sense reasoning on deductive logic, and challenges the fundamental assumptions of logic-based AI.

His reputation worldwide is out-stripped, perhaps, only by his local reputation within the Yale CS department. Since the fall of 1992 when I arrived as a graduate student, I've had the pleasure of taking classes taught by him, TA-ing courses for him, and having the occasional metaphysical chat with him over lunch. I've seen him interpret obscure student-written LISP code quicker than speed-readers scan headlines. (Rumor has it that he has a LISP parser on one of his temporal lobes.) One professor acknowledges that potential programming questions on our qualifying exams are

eliminated for being too hard if McDermott can't write out a solution in less than 15 minutes. (Graduate students are given 1-2 hours for such problems.) He has a knack for zeroing in on the holes of any theory, blasting entire logical structures with a few well-placed counter-examples. His fluency with the jargon of cognitive science, philosophy, and psychology, as well as his native computer science, impresses students and colleagues alike. If there is an AI muse, she has graced his soul.

With that kind of experience, talent, knowledge, and candor, McDermott's opinions of AI are worth hearing:

Q: What is AI?

A: It's the study of mental faculties by the use of computational models. If I build a model of a mouse, that's AI because I'm modeling a creature existing in a realistic environment. Just as biology tries to understand life, AI tries to understand mind (that's actually too grand a term; AI tries to understand control of organisms). I don't think AI is really about intelligence. The phrase is historically interesting, but if we were to do it all over again, we'd call it something else, I don't know what, ``cognodynamics'' or something.

Q: What do you think is the difference between AI and biology, or AI and cognitive science?

A: The only difference is that cognitive science is more committed to understanding existing biological cognition, as opposed to all possible cognition, which is what AI studies. So, to cite a well-worn analogy, cognitive science is to AI as **ornithology** is to aerodynamics.

Q: How, then, would you say AI is distinct from engineering?

A: I think AI researchers are always trying to make the problems they study a tad more general than an engineer is happy with. They force the problem into an area which is a little less structured, so the solutions are a little less defined, and we always have to use heuristics to solve them. You could say that the definition of the problem is part of AI. For example, with the visual tracking work that's going on here at Yale, the problem is to track something. But what's a ``thing''? What counts as a ``something''? Well, we're not sure. We look at things that people seem to be able to

track, and those are the things the human visual system picks out as trackable things, but that's not a very well-defined concept. Nonetheless, we feel that there's a real problem there, and we can show that the solutions are of great benefit in certain cases. An engineer might throw his hands up at a problem like that and say ``Bring me something a little bit more defined than that. I'm willing to print bar codes on a thing, so it's clear what it means to track them.'' In the long run, it's conceivable that there will be no difference between engineering and AI. It's hard to say.

Q: What is AI's greatest success so far?

A: Two areas that come to mind are computer vision and learning theory. They're successful in terms of achieving theoretical insight. In each case, the subdisciplines have worked out theoretical frameworks that say what we can expect, and have also developed actual working programs that do fairly well within those frameworks, although there's still much more to do. Furthermore, it's not just a matter of fads; people have written theorems that say what you can expect from learning and written practical algorithms that learn things. Similarly for vision.

Q: Where has AI failed?

A: AI's greatest disappointment is the failure to come up with a general theory of intelligence, or anything like a general theory of intelligence. In 1970, if you had asked what's the general theory of intelligence, you would have gotten answers like, ``It's all means-ends analysis,'' or ``It's all theorem proving,'' but there are very few people nowadays who say it's all anything. There's just no answer to that question. I don't see anything on the horizon that will provide that kind of theory in the future.

Q: To what extent do researchers in AI tend to bias the way they approach their work? For example, in robot navigation, we often build some kind of a map, but there are plenty of people who are able to navigate perfectly well in the world without a conscious map.

A: There's always a danger that when you use AI's approach to defining problems, you wind up defining a problem on purely introspective grounds. You always have to fight that and instead, try to abstract from your vague introspections, try to turn it into statement of the task, not the methods you would guess people are using to solve that task. You'll still get an interesting task, but you should free yourself from what you think is appropriate and not duplicate some strange idea of what humans are doing. So

if you're doing robot navigation, you don't say, ``When I do navigation for myself, a visual map pops into my mind, so I'll use a map.'' Instead you say, ``What I would like is to let it do the following thing: let it wander around a building it's never seen before, and in addition to being able to notice locations, it can also locate certain objects like bananas, books, Coke cans... After it can navigate, we can tell it what objects we're interested in and we can have it collect those.'' Now, there's a well-defined task, and it may or may not be necessary to construct a map. We have to be open to any proposal and see if anything like maps pop out.

Q: To what extent would it help to know what human brains do to solve a task?

A: I wouldn't mind knowing what people actually do, but we just have no idea what people actually do, and introspection doesn't tell us. There's at least two things you can do. One is think about how to solve the problem at all. The other is to find out what's happening in the brain. It seems to me inevitable that if you find only three methods in the entire universe on how to solve that problem, probably one of them is reflected in the brain somewhere. But you should study the task and the solutions before or simultaneously as when you examine the brain.

Q: Switching tracks: What initially got you interested in AI?

A: Well, when I was 10 years old, I saw these ads for a computer called the ``Geniac'' from the back of a comic book or something, so I asked for it for Christmas. It was a sort of plastic panel with these giant rotary switches that you wired up by screwing screws into this big rotating thing and connecting them to wires. By rotating dials, you could get it to do things. So, I diligently wired it up, and got it to do little tasks like Boolean logic and so on. Tic-Tac-Toe was the most complicated thing it could do. I remember staring at it and thinking, ``Where's the part that thinks,'' because it said that it was a computer, and it then dawned on me that there wasn't any one part that thinks. The whole thing thinks (if you're lucky). That was when I got interested in AI.

Q: Do you have any advice to aspiring AI researchers?

A: I myself would definitely have taken mathematics more seriously when I was a student. In 1970, there was this idea that computer science was a paradigm that would supplant mathematics, but it's clear now how bogus that was. I wish I had taken fewer computer science courses and more mathematics courses as a student. Save the AI for

grad school. I think if you took one AI course and one psychology course as an undergrad, that would be adequate. The computational models that psychologists study are so naive, it just boggles the mind. Neuroscience might be a good thing to take. It's helpful to know what the brain does. Depending on the subfield, you might also take engineering courses. Definitely take math.

Another piece of advice to the budding AI researchers is whenever you come across a topic that seems really exciting, turn your back on it. Anything that seems exciting to the novice is usually hot air. It's really easy in your early stages of your development as an AI person to think you're really manipulating these abstract concepts around, when it's really just fantasy. If somebody starts talking about meta-rules and you're thinking, ``Wow, this is really neat, this is like something that is self-conscious, it's able to think about it's own structure, etc.'', think again. I cringe when I hear people talk about meta-knowledge representation of something. It seems to me there are still a lot of people in AI who are living in a fantasy world, and it ultimately hurts the field. You can really only sustain the illusion that a program is actually intelligent for so long. You'll discover that you wasted a lot of time and will have to start over.

Q: When do you think such terms will cease to be just hot air? When do you think anything like human thought will become implemented?

A: I have no idea when something like that will actually happen. Or even if. The terms of the debate will change when robots become more common, as I'm sure they will. Not intelligent robots, but just robots running around doing various trivial tasks. It seems to me we'll start thinking of them as being more alive. Not literally, but they'll twinge the aliveness detectors in our brains. We'll project feelings onto them like we do to the family pet. We'll think of them as having personalities and what not, regardless of whether they actually do. Assuming that happens, the question of when they become really intelligent is very hard to answer. There are people who think their cats are very intelligent, and I assume they're wrong, but in a sense it doesn't matter whether they're right or wrong. We're not sure there's a well-defined general-purpose intelligence problem. So we don't feel any obligation to solve it. There seems to be a wide range of better-defined solvable problems, so we hope that by solving many of those, some will give insights into solving the more general problem.

Q: Do you think there are advances in other fields that might propel AI forward?

A: When I was in grad school, there was a tendency to believe AI was a paradigm competing with other paradigms. So we would say, we're not going to use Kalman filters [from engineering], we'll use AI. Nowadays, AI simply absorbs those techniques. Those techniques will continue to be of great importance. Anything at all that might be considered a part of a theory of control of an organism would be used by AI.

Q: Are you happy with your current position as an AI researcher?

A: I'm thrilled with my station in life now. I think AI has matured a lot and has developed into a much more hard-edged discipline. I'm glad just to be a part of that.