CRAIG: [00:07](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=7.15) This is Craig Smith with a new podcast about artificial intelligence. This week I talk to [Richard Sutton](https://en.wikipedia.org/wiki/Richard_S._Sutton) who literally wrote [the book on reinforcement learning](https://www.amazon.com/Reinforcement-Learning-Introduction-Adaptive-Computation/dp/0262039249/ref%3Dsr_1_1?keywords=richard+sutton&qid=1553283734&s=gateway&sr=8-1), the branch of artificial intelligence most likely to get us to the holy grail of human-level general Intelligence and beyond. Richard is one of the kindest, gentlest people I've ever interviewed - and I've interviewed a lot - a cross between Pete Seeger and Peter Higgs with maybe a little John Muir thrown in. We talked about how he came to be at the forefront of machine learning from his early days as a curious child to his single-minded search for colleagues of like mind. His temporal difference algorithm, [TD Lambda,](https://en.wikipedia.org/wiki/Temporal_difference_learning#TD-Lambda) changed the course of machine learning and has become the standard model for reward learning in the brain. It also [plays backgammon](https://en.wikipedia.org/wiki/TD-Gammon) better than humans. Richard spoke about the future of machine learning and his guarded prediction for when we might reach [the singularity](https://en.wikipedia.org/wiki/Technological_singularity). I hope you find the conversation as fascinating as I did.

CRAIG: [01:25](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=85.23) I wanted to start by asking you why you live in [Alberta](https://en.wikipedia.org/wiki/Alberta) and how such a relatively remote place became a center for cutting edge technology that's changing the world.

RICH: [01:37](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=97) The subarctic prairies of Canada. It's one of the more northern cities in Canada, actually. [Jonathan Schaeffer](https://en.wikipedia.org/wiki/Jonathan_Schaeffer). John Schaeffer is the guy who wrote the first computer program that defeated a world champion in any major game, which was in checkers. He was there before I was. I've been there 15 and a half years now. He will be humble. He'll probably blame it on me, but a lot of it has to do with him and his coming. And it was one of the first Computer Science departments in Canada, was in [Edmonton](https://en.wikipedia.org/wiki/Edmonton) at the [University of Alberta](https://www.ualberta.ca/), but how it became strong in AI originally - because I was attracted there because of its strength as well as other things.

RICH: [02:11](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=131.721) My family traveled around a lot when I was young. I was born in Toledo, Ohio and I went through a series of places where I spent like a year each. We lived in New York for a while and Pennsylvania.

RICH: [02:23](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=143.54) Then we lived in Arizona and then we lived somewhere else in Arizona. And then finally we moved to the suburbs of Chicago - Oakbrook. So I was there from seven to 17. But I think it was formative as I was very young, we kept moving around. I don't feel a real strong association to any particular place. And maybe that's one reason why it was easy for me to move to Canada.

CRAIG: [02:43](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=163.52) Well, what did your father do that kept you guys moving or your mother, If it was your mother?

RICH: [02:47](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=167.931) My father was a business person. He was an executive. So yeah, he was moving usually to the next new better position.

CRAIG: [02:55](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=175.431) What company was that?

RICH: [02:56](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=176.24) In the end he was in Interlake Steel. He did mergers and acquisitions.

CRAIG: [03:00](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=180.7) So you were not brought up in anything related to science?

RICH: [03:04](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=184.81) Both of my parents had graduate degrees, master's degrees. My mother was an English teacher and they met in college at [Swarthmore](https://www.swarthmore.edu/).

CRAIG: [03:11](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=191.98) So where did the interest in science, was that from high school or did that start at university?

RICH: [03:18](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=198.24) I was in a good high school. I was a pretty good student. I was a little more introspective, I think, than most people. So I was sort of wondering about what it is that we are and how we make sense of our world. You know, like I used to lie down on the couch in my home and stare up at the ceiling, which had a pattern sort of thing. And then your eyes would cross, you know, it seemed like it was closer. You know what I mean? And all those things, you know, make you think. It isn't like there's just a world out there and you see it. It's like something is going on and you're interpreting it and how does that happen?

RICH: [03:47](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=227.831) So I was always wondering about that 'cause it's just a natural thing to wonder about if you're an introspective kid. But I remember in grade school and, alright, what are you going to do? I think first I wanted to be an architect for some reason, then I wanted to be a teacher then I wanted to do science. And before I was out of grade school, this is in, it would be 69, you know, there's the talk of computers and you know, the [electronic brain](http://www.unz.com/print/CookJoseph-1969/). And that was exciting. And what could that mean? I get to high school, we get to use computers. And so I'm taking a course. We learned [Fortran](https://en.wikipedia.org/wiki/Fortran). And I'm sitting there saying, where's the electronic brain? You know, this thing, you have to tell every little thing and you have to put the commas in the right places and there's no brain here.

RICH: [04:31](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=271.7) In fact, it only does exactly what you tell it. There's no way this could be like a brain. And yet at the same time, you know, somehow there are brains and they must be machines. So it definitely struck me that, you know, this is impossible that this machine that only does what it's told could be a mind. And yet at the same time as I thought about it, it must be that there's some way a machine can be a mind. And so it was these two things. This impossible thing. And yet somehow it must be true. So that's what hooked me, that challenge. And all through high school, I remember I was programming an IBM machine outside of class. I think they called it [Explorers](https://www.exploring.org/). It's kind of like a modified version of Boy Scouts.

CRAIG: [05:10](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=310.391) Right. It's the next level up.

RICH: [05:12](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=312.581) Yeah, and we had access to some company and they had a big IBM machine and so I was trying to program a model of the cerebral cortex on the machine in Explorers. So it was like [neural nets](https://en.wikipedia.org/wiki/Artificial_neural_network). It was really just like neural nets.

RICH: [05:24](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=324.58) So I was into all this stuff. It seemed to me obvious that the mind was a learning machine. So I applied to colleges and I wanted to learn about this and I wrote to [Marvin Minsky](https://en.wikipedia.org/wiki/Marvin_Minsky), when I was a kid. I still have the letter that he sent me back. He sent me back that, yeah. Excellent. Can you imagine that. I asked him basically, you know, I'm really interested in AI stuff. What should I study? What should I major in? He said, it doesn't really matter much what you study. You could study math or you could study the brain or computers or whatever, but it was important that you learned math, something like that. Anyway, he sort of gave me permission to study whatever seemed important, as long as I kept my focus.

RICH: [06:05](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=365.06) I did apply to many places. I could know even then it was Stanford and MIT and Carnegie Mellon were the big three in artificial intelligence, but it was, you know, computer science. Computers were still new and they didn't have an undergraduate major in computer science so you had to take something else. And so I took psychology because that made sense to me and I went to Stanford because Stanford had a really good psychology program.

CRAIG: [06:27](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=387.47) Neuropsychology or?

RICH: [06:29](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=389.3) So, of course, psychology involves both and I took courses in both and I quickly became disenchanted with the sort of neural psychologists because it was so inconclusive and you can learn an infinite amount of stuff and they, they were still arguing inconclusively about how it works. I remember distinctly sending a term paper in to my neuropsychology class to, I think it was to [Karl Pribram](https://en.wikipedia.org/wiki/Karl_H._Pribram). Anyway, it came back marked, saying, oh no, that's not the way it works.

RICH: [06:54](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=414.5) And I thought I had a good argument that that might be the way it works, but it was just so arbitrary that someone could say what works. Anyway, I was the kind of kid that if you got, if I got a D minus, it was crushing. It wasn't really crushing but it was disappointing. And so I resolved right then that this is not fertile ground to go into the neuro side where on the other hand, the behavioral side, very rich. You could learn all kinds of good things. The level of scholarship in behavioral psychology was very high. They would think carefully about their theories. They would test them, they would do all the right control groups. They would say is there some other interpretation of these results? And they would debate it and it just seemed like a really high level - experimental behavioral psychology seemed excellent to me and I didn't know that it was bad.

RICH: [07:38](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=458.96) I didn't know that it was [disgraced and horrible](https://en.wikipedia.org/wiki/Behaviorism#Criticisms_and_limitations). I'm just a young person and I'm just saying this is cool and I still think it's cool even if it's disgraced. And I think it's just a really good example of how fads and impressions can be highly influential and not necessarily be correct. They have to be always be re evaluated. So, but behaviorism did suffer and it's gone almost extinct now. It's recovering a little bit within the biological and neuroscience parts. But behaviorist, experimental psychology has almost gone extinct. Just a few people left and it's a great shame. Fortunately, most of the basic experiments have already been done and then more is taking place within neuroscience. But I learned enormous amounts about it. I found a professor at Stanford who was emeritus, but wasn't gone. And I learned from him and I learned from the old books and that material became the basis for my contributions in artificial intelligence.

RICH: [08:44](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=524.09) So I'm known for this [temporal difference learning](https://en.wikipedia.org/wiki/Temporal_difference_learning). Temporal difference learning, it's quite clear that something like that is going on just from the behavioral studies. And unlike all the other guys in AI, I'd read all those behavioral studies and the theories that have arisen out of them. So I'm at Stanford, I'm taking psychology. Psychology is an easy major, let's face it. And I'm learning lots about that with [Piaget](https://en.wikipedia.org/wiki/Jean_Piaget) as well as the experimental things and the perceptual things and the cognitive things. But really I'm thinking I'm an AI guy and you can't major in computer science but you can take computer science courses, you can take all of the AI courses.

RICH: [09:18](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=558.451) Stanford was a world leader, one of the top three in AI at the time, and so I'm getting fully an AI education even though my major will be psychology. I'm, as much as an undergraduate can, I am getting an education in AI and I would go to the AI lab and write programs. They had the [earliest robot arms](http://infolab.stanford.edu/pub/voy/museum/pictures/display/1-Robot.htm) there and I would program them and it was great because I was allowed to program the research arms and like the first day I broke it. I thought they would kill me and kick me out forever, but they said, oh, that happens all the time and they just fixed it and they let me keep programming it. Even as an undergraduate in the summer, I'd ride my bicycle up to the [DC Power building](http://infolab.stanford.edu/pub/voy/museum/pictures/AIlab/3559DCPowersright.JPG) where the AI lab was then. And remember this is middle of the 70s and machine learning is nowhere - totally unpopular. No one is going to do anything with learning in it. So I am not quite in sync 'cause I'm thinking learning is super important and that's nowhere to be seen.

RICH: [10:12](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=612.39) But I'm getting a lot of learning from my psychology and I'm going to the library where I could read everything, read all the AI books because they were like one row and that was it. I would do that and I'd be desperately searching, you know for the learning parts and you know I found out that I was just too late because learning was really popular in the 60s and then it went off into other fields or died off in AI. There was no faculty member that was really interested in learning. I found a faculty member who was a vision guy, [Binford](https://en.wikipedia.org/wiki/Thomas_Binford), and he was the guy who said I could use the machines at the AI lab. So the founder of AI, the guy who coined the term, [John McCarthy](https://en.wikipedia.org/wiki/John_McCarthy_%28computer_scientist%29), was at Stanford. And back then what they had was a [drum kind of disk](https://en.wikipedia.org/wiki/IBM_System/360) and it's rotating around and basically as it rotates around it generates the computer screens, all the computer screens.

RICH: [11:02](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=662.04) So there'd be only like 30 I think. And this is a hard limit. You've got 30 things on the drum and so they can show 30 screens of stuff. They have more than 30 terminals, but only 30 can be active at once. And so when you go in, first you grabbed one of these 30 and then you do your stuff. But at the peak of the day they would run out. And so one time John McCarthy comes in and there are none left and so he wants to kick me off because I'm just the undergraduate. He almost succeeds. But then Binford saves the day. He says, no, no, no, he's a real person. Suck eggs, McCarthy, all people are equal. You are the creator of the field and he's an undergraduate, but he was here first so. So thank you Binford, Tom Binford. Ultimately I did my senior undergraduate thesis on a neural net model and to supervise it, I found somebody, a biologist, he was the guy who did modeling like mathematical modeling of biological phenomenon.

RICH: [11:54](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=714.19) So it's going to sort of fit. But you know, the whole thing as a way of doing AI didn't make much sense to him. And so I was again totally out on my own I guess. So I did all that. And then when I was in the library reading all those things, I was always looking for guys doing learning. I like to think in retrospect that I was looking for something that was real learning that was like reinforcement learning because reinforcement learning is an obvious idea if you study psychology. Because there are two basic kinds of learning: [Pavlovian conditioning](https://en.wikipedia.org/wiki/Classical_conditioning) and instrumental or operant conditioning. Pavlovian conditioning is for like ring the bell and then you give the dog a steak. And after a while, just from you ring the bell, he salivates showing that he anticipates the steak’s arrival. So it's a kind of prediction learning.

RICH: [12:33](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=753.67) And then there's this behavioral revealing of the prediction. The salivation reveals that the animal's predicting. For other things, you may predict them, but it's not revealed in your behavior. And so typically this kind of prediction learning has been done with things being predicted that evoke responses so that then you can see the prediction or some things related to the prediction.

RICH: [12:54](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=774.13) And then there's also control learning and control learning is called [instrumental conditioning or operant conditioning](https://en.wikipedia.org/wiki/Operant_conditioning) - at least those two names - where you're changing your behavior to cause something to happen. In Pavlovian conditioning, your salivation doesn't influence what happens. Whereas the canonical operant conditioning is, you know the rat presses a bar and then gets a food pellet. The idea of the pressing the bar's instrumental in getting the reward. So that's really the idea of reinforcement learning. It's modeled after this obvious thing that animals and people do all the time.

RICH: [13:25](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=805.21) How can that not exist as an engineering activity? It's still an oddity. I think somebody should have done it, like mathematical psychologists maybe, or maybe the engineers, or maybe the AI guys or maybe the mathematicians. Somewhere, you know, in all of the things that all the people have ever studied all over the world, surely someone else would have studied this thing that animals and people are doing, a very common-sense thing. And that's what was the mystery. So there's this mystery, but to tell the story sequentially. Last year, third year college, I come across this obscure tech report from this fellow [Harry Klopf](https://www.chessprogramming.org/A._Harry_Klopf), A. Harry Klopf, K L O P F. He has a theory that there should be neurons in a neural network that have goals, that want something. And so he's thinking outside the box and out of nowhere he's saying, well, there was neural networks before, but none of those adaptive networks that were studied earlier, none of those learning systems that were studied earlier actually wanted something.

RICH: [14:25](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=865.03) They were told what they should do, but they didn't like vary their behavior in order to get something, they didn't want anything except to drive their error to zero. And if you think about [supervised learning](https://dataconomy.com/2015/01/whats-the-difference-between-supervised-and-unsupervised-learning/), you don't affect the world. You're told the right label and then you do emit a label, your label doesn't affect the world. It's compared to the label, but it doesn't influence the next label. It influences whether you're right or wrong, but it doesn't influence the world. So then, you know, there's this question as you look back at the really early stuff before the 70s. Did they do learning where the actions influenced the world or did they not? And they would use the language of trial and error and maybe they were thinking a little bit about affecting the world. But then over time as they came to formalize what they were doing and do it more clearly, they formalized it into supervised learning and they left out the effect on the world.

RICH: [15:14](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=914.33) And so Harry Klopf was the one guy who saw this and was always complaining to himself, you know, there's something missing here. And he was the guy who saw that and [Cambridge Research Labs](https://en.wikipedia.org/wiki/Air_Force_Research_Laboratory) - that's where he was when he came to his resolution, wrote his report, then he went into the Air Force. He was a civilian researcher in the Air Force. And I'm finding his report. And so I write this guy and said, you're the only guy I can find who is talking about the kind of learning that seems important. And so he writes me back, he says, oh, that's interesting. That's what I've been saying, but no one else sees it. Maybe we should talk. So I don't know. He came to California maybe for some other reason but we arranged to meet and I met him there and that was cool. And he told me that he was funding a program at the [University of Massachusetts](https://www.massachusetts.edu/) to begin soon, or was beginning soon on this sort of thing.

RICH: [15:59](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=959.87) So as I said, he had gone into the Air Force. No one was interested in doing proper learning, learning that gets something. He told me about this thing at the University of Massachusetts and so I ended up applying to go to graduate school at the University of Massachusetts. And I say it a little bit that way because it was sort of a year early, like, you know, I met this guy, learned about this program. I thought that'd be a really cool place to go to. I think I told my mom and she says, well, why don't you go there? You know, like now! You know, because I added up my credits and I figured out that I could, if I took a couple of extra courses or something in the summer, I could finish at the end and then the very next September I could be there. And I did that. And that's what happened.

CRAIG: [16:41](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1001.84) This is for -

RICH: [16:42](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1002.421) a PhD program. So I ended up going there.

CRAIG: [16:44](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1004.78) This is at UMass.

RICH: [16:47](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1007.21) Yeah, University of Massachusetts where Harry Klopf, as now in the Air Force had gotten to a situation where he could get funding for the University of Massachusetts to do research in this area. I don't think it was his main job to fund programs in general, but he definitely set up this program to do the adaptive networks research. [Andy Barto](https://en.wikipedia.org/wiki/Andrew_Barto) and I, our task was to figure out what these ideas really were and how to think about them and how they compared to what had gone before.

CRAIG: [17:19](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1039.47) So what was the breakthrough then?

RICH: [17:22](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1042.05) The breakthrough was this was our task to figure out what he was saying. What's new because there are several ideas mixed in. There were ideas related to temporal difference learning. There were ideas related to just raw reinforcement learning, a system that wants something and tries to influence the world. But these ideas were all kind of mixed together and so as Andy Barto and I teased them apart and as we learned more and more about what has been going on in all the various potentially related fields and continued to find gaps. The breakthrough was saying, well maybe just this first part just as part of an adaptive unit, a neuron-like unit that varies its output in order to influence its reward-like signal that's coming from the world. Maybe just that. You leave out the complicated confusing parts and just do this. Like, we can't find that having already been done in the literature so maybe we should make that a thing.

RICH: [18:12](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1092.69) The name reinforcement learning in the literature - there are things that come close and that you might call reinforcement learning. We wanted to respect the existing name and the existing work rather than pretend to invent something new. There are a few antecedents so we continued to use the name. And I remember that was a real decision. It's a somewhat old sounding name, reinforcement learning. It's not really an attractive name. Maybe we should have called it, you know, something, ‘dynamic learning’ or ... So we went with that and that was the breakthrough and we wrote some papers but it was a breakthrough for us and a slow breakthrough for the field. What we were saying to the field is there is a problem that you have overlooked. Okay, there's pattern recognition, supervised learning, it would go by many names like pattern recognition and trial and error. The [perceptron](https://en.wikipedia.org/wiki/Perceptron) is supervised, all these things.

RICH: [19:02](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1142.701) So we're not giving them a better solution to pattern recognition. If you do that, you will be recognized quickly. But if you say, no, I'm not solving your problem better. I'm saying you've been working on the wrong problem, there's this different problem that is interesting and we have some algorithms for this different problem, new problem. That's the hard road. Whenever you bring in a new problem, it's harder to get people excited about it. And we're definitely going through this long period of saying this is different than supervised learning. Because they're going to say, okay, you're doing learning and you're talking about errors. We know how to do trial and error learning. It was all about making the claim and substantiating the claim that this kind of learning is different from the regular kind of learning, the supervised learning.

CRAIG: [19:41](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1181.67) Are you using neural nets at this point?

RICH: [19:44](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1184.581) Yes. So the program that Harry Klopf funded at the University of Massachusetts in 1977 was a program to study adaptive networks, which means networks of adaptive units all connected together.

RICH: [20:00](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1200.63) Maybe it's a better name than neural networks actually because the networks are not made of neurons. They're not biological. They are new units, and so it was exactly the same idea. I was also doing that kind of modeling even as an undergraduate, but it was out of fashion and Harry Klopf comes with the Air Force money, comes to the University of Massachusetts. They had several professors that were interested in in what they call brain theory, became neural networks. Brain theory and systems neuroscience. And Massachusetts was not an arbitrary choice to do this, it was the natural choice. They had a program in [cybernetics](https://en.wikipedia.org/wiki/Cybernetics), which is brain theory and AI. Brain theory was not a widely used buzzword and it still isn't, but they called it brain theory and artificial intelligence. The current wave of neural networks is the third wave. And the second wave was in the late eighties and the first wave was in the 60s. And so it was in the 60s neural networks were popular, well learning machines were really popular.

RICH: [20:56](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1256.41) And then when I was in school that was at its bottom. And so Harry Klopf, when he was saying no, we should be doing neural networks, adaptive networks, and they should have units that want things. that is totally out of sync because all learning is out of favor. So UMass, the University of Massachusetts has to be given credit that they were also interested in the somewhat unpopular topic. So maybe they were like the leading edge of the recovery. But when Harry Klopf, this odd guy from the Air Force comes to them and says, you know, I want you to work on this, I'll give you a bunch of money to work on this. And it's not military funding, it's from the Air Force, but it's totally basic research, they are going to say yes. They may not really believe in it, but they're going to say yes.

RICH: [21:39](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1299.37) So there are three professors, [Michael Arbib](https://en.wikipedia.org/wiki/Michael_A._Arbib), who is the brain theory guy, cybernetics guy, and [Bill Kilmer](https://www.umass.edu/newsoffice/article/obituary-william-l-kilmer-professor) and [Nico Spinelli](https://www.linkedin.com/in/nico-spinelli-3a124161/). So they are going to say yes, but at the same time they're not really into it. You know, it is unpopular and they're not really into it. So they find Andy Barto, Andy Barto is a professor at one of the SUNY's but he's not really happy. So he accepts this demotion from a professor to a post doc, but it's to work on this exciting problem. And he comes there and then within a year I arrive and Andy is really running the whole program. The professors, you know, they're leaving it all to him and he and I end up, you know, working on it and figuring it out.

CRAIG: [22:17](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1337.651) Did you spend the bulk of your career then in UMass?

RICH: [22:22](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1342.12) Graduate school at UMass, one year of postdoc, then I went to [GTE Laboratories](https://troy-halverson.appspot.com/GTE.htm) in Waltham, Massachusetts and was there for nine years doing fundamental research in AI and there was an AI winter and they laid lots of people off.

RICH: [22:36](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1356.34) I ended up just quitting and I go home and Andy and I started writing the book. I become a research scientist, kind of research professor at the University of Massachusetts. Sort of a distance relationship. I go there a couple of days a week cause I live a couple hours away. So I spent like four years writing the book. Part of which time I'm a research faculty at UMass. So basically living out the winter, the AI winter. And AI recovers a bit and it was that year in '99, '98 '99, and I go work at [AT&T Labs](https://about.att.com/sites/labs_research) because they're ready to do fundamental research in AI. And by then the book was out, the first edition of the book was out and I was coming back a little bit, sort of known for this reinforcement thing. There was a long period there where Andy and I were talking about reinforcement learning and trying to define what it is, trying to make progress on it. Temporal difference learning. And so you know, that's where we sort of became known as the guys that were promoting this thing and it was gradually becoming realized that it was important.

RICH: [23:35](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1415.98) Temporal difference learning in 1988, [Q learning](https://en.wikipedia.org/wiki/Q-learning) 1989 and [TD Gammon](https://cling.csd.uwo.ca/cs346a/extra/tdgammon.pdf), I think it was 1992 where he had just applied temporal difference learning and a multilayer neural network to solve backgammon at the world champion level. And so that got a lot of attention and he was using my temporal difference learning algorithm, TD Lambda. But we know we were just slowly making progress on reinforcement learning and how to explain it, how to make it work.

RICH: [24:05](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1445.211) Temporal difference learning is a perfectly natural, ordinary idea. You know, I like to say it's learning a guess from a guess. And it's like okay, I think I'm winning this game. I think I have a 0.6 chance of winning. That's my guess. And I make a move. The other guy makes a move. I make a move. The other guy makes a moves and I say, Oh wait no, I'm in trouble now. So you haven't lost the game. Now you're guessing you're might lose. Okay. So now let's say I estimate my probably of winning as 0.3. So at this one point you said it was 0.6 now you think it's 0.3, you're probably going to lose. So you could say oh that 0.6 that was a guess. Now I think that was wrong ‘cause ‘cause now I'm thinking of 0.3 and there are only two moves to go so it probably wasn't really 0.6. Either that or I made some bad moves in those two moves. Okay, so you modify your first guess based on the later guess and then your 0.3 you know you continue, you make a guess on every move.

RICH: [24:56](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1496.24) You're playing this game. You're always guessing. Are you winning or are you losing. And so each time it changes. You should say the earlier guess should be more like the later guess. That's temporal difference learning. The thing that's really cool here is no one had to tell us what the real answer was. We just waited a couple of moves and now we see how it's changed. Or wait, one move and see how it has changed. And then you keep waiting and finally you do win or you do lose and you take that last guess, maybe that last guess is pretty close to whatever the actual outcome is. But you have a final temporal difference. You have all these temporal differences along the way and the last temporal difference is the difference between the last guess and what actually happens, which grounds it all up.

RICH: [25:33](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1533.291) So you're learning guesses from other guesses. Sounds like it could be circular, it could be undefined, but the very last guess is connected to the actual outcome. Works its way back and makes your guesses correct. This is how [AlphaGo](https://en.wikipedia.org/wiki/AlphaGo) works. [AlphaZero](https://deepmind.com/blog/alphazero-shedding-new-light-grand-games-chess-shogi-and-go/) works. Learning without teachers. This is why temporal difference learning is important. Why reinforcement learning is important because you don't have to have a teacher. You do things and then it works out well or poorly. You get your reward. It doesn't tell you what you should've done. It tells you what you did do was this good. You do something and I say seven and you say seven? Okay, what does that mean? You don't know. You don't know if you could have done something else and got an eight. So it's unlike supervised learning. Supervised learning tells you what you should have said. In supervised learning, the feedback instructs you as to what you should've done. In reinforcement learning the feedback is a reward and it just evaluates what you did. What you did is a seven and you have to figure out if seven is good or bad - If there's something you could've done better. So evaluation versus instruction, is the fundamental difference. It's much easier to learn from instruction. School is a good example. Most learning is not from school, it's from life.

CRAIG: [26:49](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1609) At this point, supervised learning has been a pretty well explored…

RICH: [26:55](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1615.86) [Laughs] Just laughing cause that's what we felt at the time. There's been lots of supervised learning, lots of pattern recognition. There's been lots of systems identification. There's been lots of the supervised thing. And they're just doing curlicues and they're dotting the last i's and how much longer can this go on? Right? It's time that we do something new. We should do reinforcement learning. And that was, you know, 30 years ago or something. It's still just as true today. It still seems like supervised learning, we should be done with it by now. But it's still the hottest thing. But we're not struggling for oxygen, you know. It is the hottest thing, but there is oxygen for other things like reinforcement learning.

RICH: [27:33](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1653.671) But as people become more concerned about operating in the real world and getting beyond the constraints of labeled data, it seems like they're looking increasingly towards reinforcement learning or unsupervised learning.

RICH: [27:47](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1667.41) Reinforcement learning involves taking action. Supervised learning is not taking action because the choices, the outputs of a supervised learning system don't go to the world. They don't influence the world. They're just correct or not correct according to their equaling a correct human provided label. So I started by talking about learning that influences the world and that's what's potentially scary and also potentially powerful.

CRAIG: [28:12](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1692.49) Where does it go from here? I mean, reinforcement learning.

RICH: [28:16](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1696.72) Next step is to understand the world and the way the world works and be able to work with that. AlphaZero gets its knowledge of the world from the rules of the game. So you don't have to learn it. If we can do the same kind of planning and reasoning that AlphaZero does, but with a model of the world, which would have to be learned, not coming from the rules of the game, then you would have something more like a real AI.

CRAIG: [28:42](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1722.13) Is there a generalization?

RICH: [28:43](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1723.99) They all involve generalization. Yup. So the opposite of generalization would be that you'd have to learn about each situation distinctly. I could learn about this situation and if it exactly comes back again, I can recall what I learned. But if there's any differences, then I'm generalizing from the old situation to the new situation. So obviously you have to generalize. The latest and greatest method to generalize, it's always been generalization in the networks. Enable you to generalize. So deep learning is really the most advanced method for generalizing. You may have heard the term ‘[function approximator](https://en.wikipedia.org/wiki/Function_approximation).’ I may use the term function approximator. They are the function approximator. They approximate the function, in this case the function is a common function. Is it a decision function. If you wanted to generalize, you might want to use some kind of neural network.

CRAIG: [29:36](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1776.19) When did you move to Alberta? And why did you move to Alberta?

RICH: [29:38](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1778.831) When I was at AT&T, I was diagnosed with cancer and I was slated to die. So I wrestled with cancer for about five years.

CRAIG: [29:47](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1787.87) What kind of cancer?

RICH: [29:48](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1788.701) A melanoma. [Melanoma](https://www.medicalnewstoday.com/articles/154322.php), as you may know, it's one of the worst kinds. Very low... By the time they found it had already spread beyond the sites, was already metastasized. Once it has metastasized there's only like a 2% chance of real survival. And so we did all kinds of aggressive treatment things and I like had four miraculous recoveries and then four unfortunate recurrences and it was a big long thing. In the midst of all that, another winter started and I almost didn't care ‘cause I was already dying. So another AI winter was not really that much of a concern to me at the time.

RICH: [30:24](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1824.46) They lay off all the AT&T guys, the machine learning guys. So I'm unemployed and expecting to die of cancer, but I am having one of my miraculous remissions. So it's going on long enough, you know, how long can you go on just being set to die and going through different treatments when the cancer comes back. After a while you think, well if this, this is dragging on, I might as well try to do something again and take a job. So I applied for some jobs, but it was still kind of a winter, there weren't that many jobs and besides, I really am expected with very high probability to die from this cancer. So it's totally a surreal situation, you see. What Alberta did right is they made this opportunity for me. They made me a really nice situation to tempt me with and it was like, well, I'm probably dying but I might as well do this while I'm dying.

RICH: [31:18](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1878.76) So I say there's three reasons I went to Alberta. The position - the position was very good. It was an opportunity to be a professor, have the funding taken care of, step right into a tenured, fancy professordom, and the people, because the department was very good in Ai. Some top people like Jonathan Schaeffer I mentioned and [Rob Holte](https://www.ualberta.ca/science/about-us/contact-us/faculty-directory/robert-holte) and [Russ Greiner](https://en.wikipedia.org/wiki/Russell_Greiner) and they were bringing in some other machine learning people at the same time. They were bringing [Mike Bowling](https://www.ualberta.ca/science/about-us/contact-us/faculty-directory/michael-bowling) and [Dale Schuurmans](https://ai.google/research/people/DaleSchuurmans) and they did actually all arrive at the same time as me. So and the third P is politics because you know the US was invading Iraq in 2003 at least about the time when I was making decisions about what to do. So it all seemed very surreal. Finally I went there to accept the job and all that takes a while. And in the meantime, you know, the cancer's coming back and I'm getting more extreme treatments and by the time I set off for Alberta, momentarily in remission, and I go there and the first term, you know, and then it comes back again and it looks like I'm dying again. But then you know, miraculously the fourth time or the fifth time it works and I'm alive. One of the tumors was in my brain, in the [motor cortex](https://en.wikipedia.org/wiki/Motor_cortex) area in the white matter. So it affects my side, very similar to a stroke.

CRAIG: [32:40](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1960.13) What do you wish people understood about reinforcement learning or about your work?

RICH: [32:44](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=1964.91) The main thing to realize is it's not a bizarre artificial alien thing. Ai Is, it's really about the mind and people trying to figure out how it works. What are it's computational principles. It's a really, a very human centered thing. Maybe we will augment ourselves. Maybe we'll have better memories as we get older and we will, maybe we'll be able to remember the names of all the people we've met better. People will be augmented by Ai. That will be the center of mass. And that's what I want people to know that this activity is more than anything, it's just trying to understand what thought is and how it is that we understand the world. It's a classic humanities topic. What Plato was concerned with. You know, what is, what is a person, what is good, what does all these things mean?

RICH: [33:28](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2008.31) It's not just a technical thing.

CRAIG: [33:30](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2010.38) Is there an expectation as reinforcement learning can generalize more and more where you create a system that learns in the environment without having to label data,

RICH: [33:43](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2023.81) Yep. But that's always been a goal of a certain section of the AI community.

CRAIG: [33:48](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2028.15) Yeah. Is that one of your hopes with reinforcement learning?

RICH: [33:51](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2031.97) Oh, yeah. That's certainly, I'm in that segment of the community. Yeah.

CRAIG: [33:55](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2035.46) How far away do you think we are from creating systems that can learn on their own?

RICH: [34:01](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2041.16) Key, and it's challenging, is to phrase the question. I mean, we do have systems that can learn on their own already. AlphaZero can learn on its own and learn in a very open-ended fashion about chess and about Go.

CRAIG: [34:14](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2054.41) Right, in very constricted domains.

RICH: [34:15](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2055.801) What I would urge you to think about is what's special about those is we have the model, the model is given by the rules of the game.

RICH: [34:23](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2063.27) That domain is quite large. It's all chess positions and all Go positions. It's that we're not able to learn that model and then plan with it. That's what makes it small. It makes it narrow. So we're not quite sure what the question is - either of us - but I do have an answer, nevertheless. And my answer is stochastic. So the median is 2040 and the 25% probability is 2030. Basically it comes from the idea that by 2030 we should have the hardware capability that if we knew how to do it, we could do it, but probably we won't know how to do it yet. So give us another 10 years so the guys like me can think of the algorithms, you know, because once you have the hardware there's going to be that much more pressure because if you can get the right algorithm then you can do it, right.

RICH: [35:08](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2108.15) Maybe now even if you had the right algorithm, you couldn't really do it. But at the point when there's enough computer power to do it, there's a great incentive to find that algorithm. 50% 2040 10% chance, we'll never figure it out because it probably means we've blown ourselves up. But the median is 2040. So if you think about that, 25% by 2030, 50% 2040 and then it tails off into the future. It's really a very broad spectrum. That's not a very daring prediction, you know. It's hard for it be seriously wrong. I mean, we can reach 2040 and you need to say, well, Rich. it isn't here yet. And I'll say, well, I said it's 50/50 before and after 2040 you know, that's literally what I'm saying. So it's going to be hard for me to be proved wrong before I die. But, uh, I think that's all appropriate.

RICH: [35:54](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2154) We don't know. But I don't agree with those who think it's going to be hundreds of years. I think it will be decades. Yeah, I'd be surprised it was more than 30 years.

CRAIG: [36:01](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2161.41) And we're talking at this point about what people call [artificial general intelligence](https://en.wikipedia.org/wiki/Artificial_general_intelligence).

RICH: [36:07](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2167.11) It's exactly the question that you couldn't formulate, so let me, how would I say it? I don't like the term because AI has always intended to be general in general intelligence. Really the term artificial general intelligence is meant to be an insult to the field of AI. It's saying that the field of AI, it hasn't been doing general intelligence. They've just been doing narrow intelligence and it's a really, a little bit of a of a stinging insult because it's partly true.

RICH: [36:36](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2196.21) But anyway, I don't like the sort of snarky insult of AGI. What do, what do we mean? We mean I guess I'm good with human level intelligence as a rough statement. We'll probably be surpassing human level in some ways and not as good in other ways. It's a very rough thing, but I'm totally comfortable with that in part because it doesn't matter that it's not a point if the prediction is so spread out anyway, so exactly what it means at any point in time is not so important. But when we would roughly say as that time when we have succeeded in creating through technology systems whose intellectual abilities surpass those of current humans in roughly all ways - intellectual capabilities. So I guess that works pretty well. Maybe not. Yeah, we shouldn't emphasize the 'all.' But purely intellectual activities, surpass those of current humans. That's all right.

CRAIG: [37:32](https://www.temi.com/editor/t/96tRIVYgwEqXG8kt88zs_etBbWGPlRqQoUC_2Lx7gCTH-HeQeug0BHfKVBoFM7rS4Kv5KroYrc92AKsDky9If-_aoqk?loadFrom=DocumentDeeplink&ts=2252.8) That's all for this week. I want to thank Richard for his precious time. For those of you who want to go into greater depth about the things we talked about today, you can find a transcript of this show in the program notes along with a link to our [Eye on AI newsletters](https://www.eye-on.ai/subscribe-page). Let us know whether you find the podcast interesting or useful and whether you have any suggestions about how we can improve. The singularity may not be near, but AI is about to change your world. So pay attention.