EPISODE 94

[INTRODUCTION]

[0:00:10.8] SC: Hello and welcome to another episode of TWIML Talk, the podcast where I interview interesting people doing interesting things in machine learning and artificial intelligence. I'm your host, Sam Charrington.

Just a couple of quick announcements today related to the TWIML online meetup. First, the video from our December meetup has been posted and it's now available on our YouTube channel and at twimlai.com/meetup. It was a great meetup. If you missed it, you'll definitely want to check it out.

You definitely don't want to miss our next meetup either. On Tuesday, January 16th at 3:00 Pacific, we'll be joined by Microsoft researcher, Timnit Gebru, who will be presenting her paper Using Deep Learning and Google Street View to Estimate the Demographic Makeup of Neighborhoods Across the United States, which has received national media attention for some of its findings. Timnit will be digging into those results, as well as the pipelines she used to identify 22 million cars and 50 million Google street view images.

I'm anticipating a very lively discussion segment as well to kick off the session. Make sure to bring your AI resolutions and predictions for 2018. For links to the paper or to join the meetup group, visit twimlai.com/meetup.

All right, on to today's show. In this episode, we hear from Kenneth Stanley, Professor in the Department of Computer Science at the University of Central Florida, and Senior Research Scientist at Uber AI Labs. Kenneth studied under TWIML #47 guest, Risto Mikkulainen at UT Austin after Geometric Intelligence, the company he co-founded with Gary Marcus and others, was acquired in late 2016.

Kenneth's research focus in neuro-evolution, which applies the idea of genetic algorithms to the challenge of evolving neural network architectures. In this conversation, we discussed the neuro-evolution of augmenting topologies or NEAT paper that Kenneth authored along with Risto, which won the 2017 International Society for Artificial Life's Award for Outstanding Paper of the Decade 2002 to 2012.

We also cover some of the extensions to that approach he's created since, including hyper NEAT, which can efficiently evolve very large neural networks with connectivity patterns that look more like those of a human brain, and that are generally much larger than what prior approach is to neural learning could produce, as well as novelty search, an approach that unlike most evolutionary algorithms has no defined objective, or rather simply searches for novel behaviors.

We also cover concepts like complexification and deception, biology versus computation and some of his other work including his book and Nero, a video game complete with real-time neuro-evolution. This is a meaty nerd alert interview that I think you'll really enjoy.

Now, on to the show.

[INTERVIEW]

[0:03:25.8] SC: All right, everyone. I am on the line with Kenneth Stanley. Kenneth is a professor in the Department of Computer Science at the University of Central Florida, as well as a Senior Research Scientist at Uber AI Labs. Kenneth, welcome to This Week in Machine Learning and AI.

[0:03:39.7] KS: Thanks very much. Really happy to be here.

[0:03:40.1] SC: Fantastic. Why don't we get started by having you tell us a little bit about your background?

[0:03:45.9] KS: Sure. I've been interested in artificial intelligence since I was a little kid, maybe around 8 years old. Went on to major in computer science because of that and carried that interest into graduate school, where I was at the University of Texas at Austin, where I did my PhD. There, I became interested in particularly neural networks, artificial neural networks which are now the basis of deep learning, which everybody is talking about. Also what's called evolutionary computation, which means Darwinian type of principles being applied inside of computer algorithms.

The intersection of those two things is what's called today neuro-evolution, which means like evolving neural networks or evolving brain. You could think of it as in a computer. I guess, my particular interest is just how brains evolve. These amazing astronomically complex things that are in our heads. I was always fascinated by how an unguided process seemingly an unintelligent process like evolution could just produce something so astronomically complex and amazing as our own brains. As a neuro-evolution researcher, I've been trying to figure out how you can actually make algorithms that would evolve something of similar scale and complexity.

[0:04:59.8] SC: Was there anything in particular that you came across at the age of 8 or so that got you interested in AI?

[0:05:05.4] KS: Yeah, yeah. At the age of 8, that's when my family bought a computer. It was a Commodore 64. It was also, my parents put me in a programming class and that was on a TRS80 with a very old computer system.

[0:05:20.6] SC: Trash 80.

[0:05:22.5] KS: Yeah, exactly. I guess, for some reason like as a little kid it just really made an impression on me that I could tell the computer to do anything. I had this feeling, there was like infinite freedom in the things that I could get the computer to do. If only I could just figure out how to tell it what I wanted, and I felt like if I could just tell it how to have a conversation with me, then it would basically be my friend or talk to me.

I was really, really interested in just getting the computer to have a conversation with me, like a casual conversation like, "How are you doing? What's your name?" That kind of thing. At first, I would rate really simply programs in basic; the basic computer language that would be little conversations like this. I'd say, "What's your name?" I'd say, "Ken," like basically in typing and it would say, "Hi, Ken." I was very impressed that we can have this kind of conversation and then I got it to do that.

I quickly hit a wall where I couldn't get it to really do anything interesting. It's just a very stock scripted thing. At the time, like around age 8 I thought there is some way to do this that I just need to read a book or something, like there's something that would just tell me how to get it to have a real conversation with me.

I didn't realize that this is one of the greatest problems facing humankind. We have to get a computer actually to be intelligent, like a real person. It took me a while actually for it to strike me. It is actually like an extremely hard problem and there's not just some manual you can read that you can get the computer to do that.

Probably within a couple of years I realized this is like a huge problem and I was really interested and hooked. Like, "Wow, this is actually hard. There's got to be a way to do this." I guess, I would just stay captivated by that problem forever. I guess, I changed and shifted a bit in my interest, because if you look at that and you look at it from the lens of today's subfields of artificial intelligence, you'd probably call that natural language processing or something like that. I shifted away from that over time to more lower level stuff, like control, like neural stuff. That was what initially hooked me into it and got me interested in AI.

[0:07:27.7] SC: Interesting. You mentioned that you studied at UT Austin. I did an interview with Risto Mikkulainen, did you study with him there?

[0:07:36.3] KS: Yeah. I guess, it's just a coincidence that Risto is my adviser. He was my adviser during the PhD, so I worked with him for years there. Yeah.

[0:07:44.6] SC: Awesome. Awesome. Can you tell me a little bit about your primary research focus?

[0:07:50.2] KS: Sure. My primary research focus is in an area called neuro-evolution. It's an area that is probably less well known in the general public, like you hear tons of stuff about deep learning today, but you don't hear so much about neuro-evolution. It's certainly related to deep learning, because both of them are about in effect neural networks. But neuro-evolution has this twist which is that we're interested in neural networks, which are for those who don't know basically these rough abstractions of what happens in brains. The word neuro comes from neurons and neurons are in our brains.

Neuro relationship roughly, it motivated or inspired by brains in nature. Although they're not at all accurate models of them. But then in neuro-evolution, we're combining that with evolutionary principles, which really means breathing, like if you think about it like it's like, if you had a neural network that does something good, like say drives a robot and makes it able to do a task, like say walk, like it gets your biped robot to walk. Then neuro-evolution is you're breeding those brains. You're saying, "Okay, I have a bunch of brains. These are artificial brains." We'll call them neural networks though because artificial brains exaggerates how cool they are to see, the artificial neural networks.

We would then look at how well did they get the robot to walk, like a whole bunch of them and they call that a population. We'd choose the ones that do better. Some will do worse and some will do better. The ones who do it better will have children, which basically means new neural networks should be born as offspring of the old ones that we chose, or we call that selection. We selected those.

Our hope is that the offspring of those better ones will sometimes be even better than their parents. If we keep on playing this game, which is just breeding, so it's not hard to understand. Some of the AI are complex and hard to understand at first, but intuitively this is easy because this is just like breeding horses or breeding dogs. You just choose the ones that are better and with respect to whatever criteria you have, then just breed them and hope that things get better over time.

Neuro-evolution is basically about breeding these artificial things, rather than real organisms, which is like these artificial neural networks. Thereby getting them to get better over generations. What is interesting about it to me is that – a simple concept and principle, at least like the initial element and it gives quite simple just in terms of breeding. Under the hood, there is real mysteries here. Because this is really the process that produced you and me and the high level of intelligence that we have going all the way back to singled cell organisms.

It's quite amazing to believe that there is some path through that space just through breeding, that can lead to something like us from something so humble and simple. To get algorithms to do that is an enormous challenge and not fully understood right now. That's where the research comes in in the field.

[0:10:50.2] SC: Interesting. Interesting. Then you're also again a senior research scientist at Uber AI Labs. What can you tell us about Uber AI Labs and how that came about and what the charter is there?

[0:10:56.4] KS: Right. There was no Uber AI Labs around nine months ago, but I was one of the co-founders of a startup company called Geometric Intelligence. My co-founders included

Gary Marcus, Zoubin Gharamani and Doug Bemis. Some of them really quite well-known and a very respected researchers themselves.

We were doing in Geometric Intelligence, proprietary machine learning research and developing new technologies and building a team that we were hoping to be a world-class research team. What happened was that Uber acquired us nine months ago in December. When Uber acquired us, they had partly one of their aspirations was to start an AI lab, like a real research lab in industry that researches the cutting-edge of artificial intelligence.

Uber believes and believed at the time that artificial intelligence is a critical competitive component of the industry, where Uber needs to be staying at the cutting-edge. Uber has and had before a lot of competence already in machine learning. It's not like there was nobody here. There were plenty of people here who were very qualified in the field. They didn't have something that was really a fundamental research lab and where they started just really pushing on the cutting-edge of AI itself, as opposed to just applying it to internal problems.

For example, Uber has a team focused already that was focused on autonomous driving. They already had that in place, but that's an applied aspect of artificial intelligence. The AI lab that was founded off of the company that we started which we founded, was really intended to be focused more in advancing the algorithms themselves.

What Uber got was basically all at once, like all of these researchers who had this capacity to push forward the field of AI. You can think about it roughly an analogy with similar types of research labs at big tech companies, maybe something like DeepMind, which was originally acquired by Google or something like Facebook AI research or something like Google Brain at Google.

There's some rough analogy there between us and them. Then we're much smaller there, because we're newer, but we have the similar mandates in terms of researching the cutting-edge of AI. I should say that actually we're going to – we are going to engage with the outside and the academic community, so we'll be hearing from Uber AI Labs. We're going to be publishing. We understand that we cannot be a successful AI lab if we are not engaged with the outside world. We will be publicizing and publishing some of our work, so people can see what we're doing, and so that we can communicate with other researchers and scientists across the world.

[0:13:43.0] SC: Okay. Great. Great. Can you talk a little bit about the intersection between your work and evolutionary AI and the things that Uber is doing around self-driving cars?

[0:13:54.5] KS: Yes. I can't get into specifics about what Uber is doing with their self-driving cars for obvious reasons, but I can say that Uber AI Labs is diverse. I mean, that was one of the original inspirations behind Geometric Intelligence of the predecessor of AI Labs was to have a diverse group that isn't just in one particular fad, which you might say deep learning is, although it's obviously an important one that's making a lot of important contributions.

Our philosophy was that we need to not have all our eggs in one basket. Uber AI Labs itself is like that too, and we have a lot of diversity in terms of the expertise and areas that we cover. Among those, we clearly are world-class in neuro-evolution, which is the field that I just described where I focus at most at my career.

This is a particular direction within AI and machine learning that offers some unique insights and angles on certain types of problems that other areas might have a different take on. In terms of autonomous driving, it's clear that the idea of the evolution of complexity and the how really high-level intelligence can be evolved in terms of complex, large deep artificial neural networks has a connection in principle to how you could get a really sophisticated controller for a vehicle or something like that.

The insights of the field of neuro-evolution both directly, which means like using neuro-evolution itself as an algorithm and indirectly in terms of insights that we gain as a side effect of doing experiments in that area can impact how we would create algorithms that might control things like autonomous vehicles. I should also note that it's not the case that the only application, or even necessarily the main application of AI at Uber is in that area.

I mean, Uber has AI problems across the gamete of all of their business components, so there is a lot of different applications that are under consideration when it comes to AI Labs and what AI Labs does.

[0:16:05.5] SC: Sure. Can you talk a little bit about how your research focus compares and contrasts with what Risto is doing down at UT Austin?

[0:16:14.5] KS: Yeah, sure. I mean, actually there's a lot of overlap, because I mean, I'm his advisee. I've taken a lot of original teachings that he gave me as a basis of my career and obviously collaborated with him for years to publish some of the – in the end, it turned out to be some of the seminal papers in the area both together.

I think we're not actually so different in terms of the fields that we're interested in. Where we may differ is more just in what particular algorithms that we contributed to inventing since we parted ways when I basically graduated with the PhD. He's focused on his own set of innovations and I've focused on my own. There is some divergence there, but we really ultimately tend to be very close. Because when I've invented new things and I was like – I'm still at the University of Central Florida. As a professor, Risto would sometimes build on those things and vice versa. We're very intertwined. It's not a surprise since we started out in the same area.

[0:17:25.3] SC: Absolutely. Absolutely. Folks that are interested in maybe some of the background on – you talked about the breeding process that are really high-level. Risto and I spent quite a bit of time digging into that into more detail. Folks that are interested in that might want to refer back to that podcast.

Since you've graduated and now that you're driving your own research agenda, what are some of the specific algorithms that you've published research on and how do they build on the core ideas of genetic or evolutionary computing or algorithms?

[0:18:06.9] KS: Yeah, sure. So neuro-evolution, which is this idea of evolving neural networks; one of the interesting things is that when – at least for me, what I find really interesting is not just optimization. A lot of people in machine learning think in terms of optimization, which means just like how do you get the structure to get better and better and better with respect to a task?

I'm also interested in what you might call complexification, which means how do we get increasing complexity? The thing that really fascinates me is like how in nature things got more complex, like insanely more complex? Not just a little bit of incremental increases in complexity, but from a single celled organism to something that has in our brain a hundred trillion connections among a hundred billion cells approximately, or a hundred billion neurons.

That's just amazing to me that some kind of unguided process could build something like that. This is not something that was engineered. I always have my eye on like what is it that allows really high-level astronomical levels of complexity to emerge from this kind of process, this automated process.

The interesting thing in neuro-evolution is that every time it would seem like we have an advanced, where we figure out something about, how do you get increasing complexity happen inside of an algorithm? We've made some advances in, including the first thing that I did in grad school, which was this algorithm called NEAT, or Neuro-Evolution of Augmenting Topologies, which I did with Risto, which was basically an algorithm about how can we have the neural networks that are evolving in the computer increasing complexity over the course of the algorithm running in the computer.

It was because I had this real fascination with increasing complexity that led to us introducing this algorithm that increases complexity. But then what's interesting is that every time we make an advance like that, it uncovers some deeper underlying question. Because it turns out that the explanation for why it was possible to get from one cell to trillions is really, really subtle and nuanced and complicated.

[0:20:09.8] SC: When you say that, are you speaking biologically or from a – in a computational context?

[0:20:17.2] KS: Right. Good question. Actually, those things constantly get intertwined in my mind. Whether I'm thinking biologically or computationally, because the way I look at it is like the biology and computation aren't really necessarily different things. In effect, if you read a biology textbook you feel like you're reading about biology. In effect, it's also about computers, because – or at least algorithms. Because you're talking about a principled process that basically follows some certain kinds of rules.

[0:20:45.6] SC: This is analog computers that we really don't understand very well.

[0:20:49.1] KS: Yeah. You could think of the universe as a big analog computer we don't really understand. I mean, evolution is a very algorithmic thing. You're talking about there are individuals and those individuals reproduce, and then the thing that – who gets to reproduce is based on a formula, which is obviously complicated, but basically some individual reproduce, some don't.

This can be formalized as basically like a program. You could imagine writing the rules of the system. This is what inspired the field of evolution or computation. I mean, people saw the theories of evolution and biology and thought like, "You know what? This is actually not that hard to write down as a program and actually make evolution happen artificially inside of a computer."

It turned out though that if you just read the textbook and then learn these principles that sound like good explanatory principles for how evolution works. Like if you read a biology text it was like, "They know how it worked." That's an explanation.

It turns out that explaining something is easier than actually implementing it, which is basically something that we found across the field of artificial intelligence. You can read about – you can read a neuro-science textbook that says this is how brains work. Of course, biologist would acknowledge you don't know everything, but this is what we understand now and it's a comprehensive explanation. It's far, far away from telling you how to actually build the brains. You don't know how to build a brain just because we have some understanding of how brains work.

It's the same with evolution. We don't know how to build a true evolutionary system at this scale and magnitude of what happens on earth. Even though we know a lot of the details about what goes on. The missing details, like the gap between what we understand and what we can actually build, that's where the research is and that's where a lot of fascinating insights occur.

To me, I think that to some extent, when we make advances in artificial intelligence, we're actually learning something about biology in a sense, because we're realizing that the gaps in our knowledge like what we didn't understand are actually filled by something that we didn't expect, or that wasn't in the textbook about how things work.

It's true that sometimes we may be doing things that are not actually the same as biology, but at least they're revealing gaps in our knowledge of biology. In some sense, if we actually knew everything about how things works, then we could just program it in, but we clearly don't.

It's like, I think AI has a higher bar in a way than biology. Biology, you can explain something or statistically analyze it, but in AI you actually have to build it, which is much, much harder. It forces us to grapple with the problems of the gaps in our knowledge in biology. Some people in AI would just say not like that way of looking at things, because some people in AI don't care about the biology and they just want to build intelligent things and they don't really care. Do these things correspond or not with biology? That's not the goal. The goal is just to build intelligent things.

We aren't adherent to biology or not. I tend to be more biologically inspired, but I also agree that I don't really, let's say ultimately care whether what I build is exactly the way it works in biology or not, but I just find it interesting and inspiring that biology has achieved things that are just so amazing. Like human level intelligence. I find it fascinating that we just don't know how. Trying to probe those gaps in my understanding, I find leads to over and over again really deep insights in artificial intelligence, because it's like we suddenly realize, "Oh, wait a second. Actually there's an explanation here which is much different than what we thought it might be."

After graduate school, there were a succession of those that I went through; we would realize that there's something missing still after like for example the NEAT algorithm, which actually became the most used algorithm in this niche field of neuro-evolution. We realized there is limitations on what NEAT can ever do.

Basically, how can we get around those limitations? How did nature get around those limitations? One example is that in NEAT, there is artificial DNA which encodes the neural network. They have to do evolution, so we have an artificial DNA and we call it genome. It would have one gene per connection in this brain that's evolving. This is clearly not going to scale, even though this brain can keep expanding, but if you wanted to get a hundred trillion connections, which is what we have in our brain right now in biology, we would need a hundred trillion genes in NEAT.

There is no way that's ever going to happen. A hundred trillion genes is just astronomically insanely large. For example, our genome in biology only has 30,000 genes or three billion base pairs, another way of thinking about it.

We had to invent new algorithms and this is after grad school and after NEAT that could encode much, much larger structures. We called these indirecting codings. This led to something called hyper-NEAT eventually, which is a new kind of genetic encoding that is much more compact than the original NEAT. Hyper-NEAT was something that I did after I left UT Austin, and so where I did that independently off Risto and led to the ability to evolve much bigger in effect neural networks.

Then I think one of the biggest things probably that has had a lot of impact in the field after that was something called novelty search, which is a result of discovering that in some cases, the best way to get something in a search process, and evolution is a kind of a search process, like you're searching through a space of possibilities is to not be trying to get it. This was a really counter-intuitive and paradoxical insight, but really important I think for realizing how things are achieved.

In other words, if you say that you're trying to breed for something. Say we want to get human-level intelligence, then that actually may doom you from the start. Sometimes, the only way to get to something is to not be trying to get it. This is a bitter pill to swallow, but something that we realized.

[0:26:43.9] SC: What is the mechanism of trying that keeps you from being able to get it?

[0:26:45.1] KS: Yeah. The mechanism there is something called deception. Actually, this is something that applies way, way outside just neuro-evolution. This is a general principle for everything in life.

[0:26:58.6] SC: Is that deception?

[0:26:59.1] KS: It's called deception, yeah. It's basically the situation when if you are observing that things are getting better. It's like you have some metric for what it means to be doing well, like a performance metric. Let's say how well are you able to walk? You have some metric that says, "Well, how well am I walking?"

Normally if I was trying to get something to walk, I would select things, meaning I would breed things that are apparently better at walking compared to their predecessors, and I call that their fitness. That's what I mean by trying. I keep on intentionally picking things that seem to be better. This is a very intuitive idea, like everybody for a long time felt that this is obviously the way to get things to evolve is to pick things that are better.

It turns out that if you're in a deceptive situation, which it turns out unfortunately you often are in, that you can be moving in the wrong direction, even though your metrics for performance is going up. That's because the world is really, really complicated. It can appear that you're improving in some way when you're actually not.

For example, when it comes to walking, lunging forward like a maniac and falling down like a few feet from where you started may appear to actually be an improvement in your ability to travel, because basically you're getting farther than your predecessors by throwing yourself on your head 5 feet in front of you.

This is actually not a good stepping stone towards really good walking behavior. In fact, a good stepping stone might be discovering the concept of oscillation. That's what your legs actually do, they oscillate when you walk. It could be that when you initially discover isolation, you fall in your face. It actually looks like you're not improving. But because your metric is basically how far did you go, it causes you to basically be blind to the underlying discovery that's actually essential to making the progress that you deem to make in the long-term. This problem of deception is just like universal across all kinds of endeavors. Not just neuro-evolution. It's like –

[0:28:55.8] SC: This is analogous to almost like a local maxima issue?

[0:29:01.3] KS: Yeah. I mean, it's basically the same thing. It's related to local maxima, or local optima, or premature convergence and sometimes people would call it to getting stuck on a local optimum. I think that the insight that we have that's different from just saying, "Okay, well we just rediscovered local optima, because we already knew about local optima."

[0:29:20.0] SC: Right. Exactly.

[0:29:21.0] KS: It's just how utterly profound the problem is, that you cannot just – I mean, people think there's ways of getting around local optima. There is tricks; we have diversity, we have randomness, stochasticity, there are things you can do to jiggle things around a little so we don't just get stuck on a peak, which is what we think of local optimas, like getting stuck on a peak in a big space.

That's just not going to cut it in certain types of problems, because they are just so absolutely complex that almost no matter what you do, deception is going to kill you. We showed this when we introduced this algorithm called novelty search. That in some problems, that it was shockingly terrible what deception could do to you in these spaces.

What was profound was that we showed that in certain problems like these where deception is a really big problem, and I would claim that deception is a really big problem in like almost any interesting problem. I can demonstrate that later if we want to get into it.

When it is a serious problem, then we show that with this novelty search algorithm that we introduced, which was basically not trying to solve a problem, but rather it was just driven by, selecting things are more novel. Not things are better, but just more novel that this would actually be better at solving a problem that was deceptive, than an algorithm that was actually explicitly being driven by selecting things that were better.

The lesson it showed is it can be better sometimes to not be trying to solve the problem than to actually try to solve the problem in terms of getting a better solution. This is obviously really counter-intuitive and paradoxical and upsetting maybe even, because it's embarrassing in a way for anybody who is saying, “Okay, I've got this really good optimization algorithm.” To lose to an algorithm doesn't even know what kind of problem it's trying to solve. That's what novelty search is. Novelty search is a divergent search algorithm. Basically, it's just trying to find things that are different than what it's found before.

[0:31:20.0] SC: It sounds a little bit like explore, exploit where your explore is optimizing for newness.

[0:31:22.5] KS: Yeah. Yeah, it is related to this exploration-exploitation dichotomy that a lot of people talk about in machine learning. It's also different, I think. There is an additional element of insight here beyond which is really important. Which is that when we think of exploitation versus expiration, like often we think of exploitation as following some gradient, which means information towards something that we are trying to get to.

In other words, we're using information to move in a direction that's intelligent. Interestingly, exploration we tend to think of as random moves are ignoring the informed gradient. It's like, let's just go somewhere and see what happens. That's what we think of exploration. But with novelty search though, there is a principled exploration that is not random, that actually exploration is something that's also very informed.

In a novelty search case, you're informed by where you've been, because novelty is basically a comparison between where I am and where I've been before. It's anything but random. It's a very informed gradient. It's just that it's the gradient of novelty instead of the gradient of the objective.

This is actually a very information-rich gradient, because when you think about it you know a lot about where you've been. In fact, you know more about where you've been than you know about where you're trying to go, because the whole problem with where you're trying to go is you don't know about it. Otherwise, you would just go there.

Novelty is actually more informed I'd say than the objective gradient. For this reason, it's an extremely interesting gradient to follow, like the gradient of novelty, because you're being pushed away from where you've been before and it turns out that you will be inevitably pushed towards higher complexity. It's really tied into this idea of increasing complexity.

Because if you think about it like, as soon as you exhaust all the simple things you can do in the world, like the only choice you have if you want to continue to create novelty is to do something more complex. Ultimately, there is an inevitability that with novelty search that you're going to be pushed towards increasing complexity.

I think of it as almost like an information accumulator. In order to continue to do novel things in the world, you have to accumulate information about the world. For example, you could imagine if you were trapped in a room and I told you to just do novel stuff. For a while, you could just run around randomly and you'd bump into walls and everything you do would be novel.

Eventually, you'd bump into all the walls in the room. At some point, you're going to have to learn how to not bump into walls. When you do that, you're going to have to learn what a wall and how to sense a wall and how to navigate walls. Eventually, you'll have to learn how to open a door, because you have to get out of the room eventually to do something new. Eventually, you're going to have to get off planet earth and go to Mars.

Clearly doing that requires learning extremely deep and complicated facets of how the universe works, like physics. You're going to be forced to become an expert on the domain where you find yourself if you're going to be pushed towards doing more and more novel things.

Novelty search actually is a very deep and interesting kind of a process. That's why sometimes, it alone will do better than actually trying to solve the problem you're trying to solve. If you think like back evolutionarily, if you think about how could we get to human intelligence from a single cell? It would be crazy to do selection based on the intelligence of single-celled organisms.

We would start up by applying IQ tests to single-celled organisms. That would just kill the population, because none of them are intelligent at all. It's funny, but in a sense, the reason that we got to where we are today is because we were not trying to get there. If we had started out where selection was based on intelligence, then everything would've died, or we would've gone nowhere and we wouldn't have gotten to where we are today.

We see this issue of deception come up over and over again. It turns out that there was this turning point long ago, eons ago where symmetry, bilateral symmetry was discovered. These are our ancestors. There is this bilateral symmetric flatworms. There's no indication that it's had anything to do with being more intelligent, but actually it does in some kind of really, really long-term sense, like that was an important discover that led ultimately or stepping stone that leads ultimately to a human level intelligence.

You wouldn't be able to predict that on the basis of doing an IQ test. Yet, we needed to lock that in. In some sense we could recognize that was interesting from a novelty perspective, because it was a very new innovation. We cannot recognize it from a performance perspective, because at that long, long ago point in time, it's not an indicator at all from the point of view of performance. Like if the ultimate indicator is intelligence.

This is another kind of example of deception and why many things are not going to be possible to discover if we just set them as a goal and just select based on those things. This is a principle not just for evolution, but for life too. There are many inventions that would not have been invented if they had been our goal to invent them, which is again the paradox coming up.

Computers for example were – the first computers were based on vacuum tubes. The people who invented vacuum tubes were not trying to invent computers. If you had gone back to the 1800s and told all the researchers working on vacuum tubes, who were interested in electricity that like, actually there's something more interesting like a computer and maybe you should just invent that. Forget this boring vacuum tube stuff. You would neither have vacuum tubes nor computers.

Once again, we needed people to be exploring very diverse ideas without having their eyes on the prize, if you think of the price as like a computer, in order to eventually get the prize. There's a paradox there.

This concept is so general and connected to this novelty search idea that we wrote this whole book about it called *Why Greatness Cannot Be Planned.* After a long time researching novelty search and a long time for me talking in various forums and venues about novelty search. I realized that the principles are really general about this paradox, this is what I call the objective paradox. That it's actually relevant to all society, like how we run our institutions. We give money to people based on them making progress with respect to an objective, like this is what granting agencies do in the sciences.

It's actually not principled in the long run. There are other processes that need to be recognized and respected if we really want to be able to achieve really, really ambitious ends. That's why we wrote this book basically to introduce these principles of deception and divergent search and the objective paradox to the general public. We were hoping that maybe this would actually provoke a discussion of these things in a larger sense, because of the fact that it affects many of the attempts in innovation that we as a society are engaged in. It turned out as really broad implications across culture and society.

[0:37:50.4] SC: Interesting. Then one of the papers that I noticed is one called Galactic Arms Race. Is that an extension of this work, or is that a different direction?

[0:37:51.6] KS: It's related. Yeah, it's related. As we started to understand this idea of – we call it sometimes divergent search; searches that are not aimed at a particular goal, but rather which are divergent to the space of what's possible.

They're searches that show you all the cool stuff that you could find, like not just one thing. Evolution on earth is like that. There's not like one thing it's trying to do. It wasn't trying to get human-level intelligence. It illuminated all of the possible cool stuff that's out there in nature, like all of the diversity of nature.

We started to realize, these algorithms are really cool that do stuff like that, perhaps for applications in the real world. In the Galactic Arms Race, the applications of video game. Our idea there was maybe we could put one of these divergent search algorithms in a video game so it would generate the content in the game, and you get more and more cool content just flowing into the game from nowhere. No human has to actually design or invent it.

In the case of Galactic Arms Race, it was the weapons of the ships that you fly. I think people are familiar in video games, like with playing games where you have to pick up new types of lasers, or weapons, or guns, or something like that. We said, "Let's let evolution invent the weapons," but with a novelty search, like process where it's not aiming for the optimal weapon. It's just diverging through the space with weapons. With some information about how humans are actually using them.

It's informed by the humans in the game and in real-time inventing new weapons for the humans to try. There's an interaction, we call interactive evolution between what humans do and what evolution does. It causes all these cool weapons to be invented. Things that I don't – never seen any other game that which invented by the computer itself.

It's kind of, I think a really nice exposition of the potential of divergent search or novelty searches to create open worlds where things are just continually generated. Sometimes we call this open-ended evolution, that are interesting and hopefully without end.

[0:39:53.9] SC: What's an example of a type of weapon that was invented in this game?

[0:39:55.9] KS: Okay, yeah. There's a couple of good ones. It is funny, we started naming these things after the fact, because they don't actually have names, because they're invented by the computer. One we call the tunnel maker, which would basically generate two streams of particles – these are all particle weapons that would very slowly shoot on the left and right side of your spaceship.

Basically, it created a protective tunnel that you could fly through. Then in the middle of that tunnel, there was another faster stream that was actually used for shooting things. You would be creating basically a shield that would shoot out from your sides that you could then fly through.

There was another one that we call the lasso, which just looked like – it looked like a cowboy's lasso. It just shot out and created the spiral around the enemy and then closed in on it.

[0:40:40.4] SC: Interesting.

[0:40:41.3] KS: It was really surprising that this thing was invented. It was interesting, because I actually – it's not a great weapon in an objective sense, like the lasso one. Because I think, it's much better probably just to shoot straight at something and kill it. The players loved it, because of the aesthetics. It's just so interesting and fun to have the lasso weapon and to show off, because it was a multiplayer game, so people could see each other's lassos. That became popular and the game just went with it.

The game didn't say, "This is objectively worse, or objectively better." It just saw that people were interested in lasso, so it created more of the lassos and diverse lassos. We had all these lasso weapons proliferated in the world, because people like them, whether they're optimal in some objective sense of not.

[0:41:20.4] SC: Is there an argument that says that the issues around – that you identified in novelty search and getting led down the wrong path. The example I guess you gave us with a robot trying to learn how to walk and using a motion that doesn't lead it towards walking and eventually lets it fall on its face. I guess, the thought's are – can all those be boiled down to just not being able to express enough sophistication in our objective function, or not being able to express our objective function in the right timeframe, or something like that?

[0:42:00.3] KS: Yeah. Actually, there's an element of truth to that view that if we knew enough about the world, we could just write the objective function to take into account how the world actually works. The problem is that in practice, that's just impossible, because you ultimately will have to know every single thing about all the stepping stones that you would have to go through to write the objective function to take that into account.

Say there is like a million steps between here and a human-level AI. Obviously, if I wrote a fitness function where your score is literally how far you are along that path, then of course this is like the ideal objective function is going to work out fine. The whole problem that we're facing just begs the question of how are we going to figure out what those stepping stones are, so we're back to square one again.

In practice, like you're probably not going to be able to do that in even like a relatively simple problem, because the whole problem with search is we don't know the stepping stones. If we did, we wouldn't be doing search, because we would just build the thing because we would know all the steps in terms of how to get it.

This paradox is basically unavoidable. If the problem is not interesting then we do know the stepping stones that we don't need to do these things, but the problem is uninteresting. If the problem is interesting, it's interesting because we don't know the stepping stone. That's what it makes it an interesting problem.

Almost any interesting problem is going to be confronting this paradox. Now that doesn't mean that there aren't some cases where search will work. Obviously it will with an objective sometimes. There's no doubt about it. In fact, deep learning has exposed that in really high-dimensional spaces between spaces of many, many parameters, like many weights in a neural network, that there's less deception than we thought.

This has been a surprise for everybody including me. Sometimes we still can just push brute force to the objective function, because high-dimensional spaces have some very odd properties and succeed at solving problems.

We shouldn't conclude from what I'm saying that our objectives are completely useless. They do work in some cases, but I think that it's still the case that in very, very complex problems we are going to be facing deception. We are not going to know how to write the correct objective function to go through all those stepping stones, which are basically reflecting eons of progress to get to some of these really ambitious ends that we have.

It's an element. It's not like everything should be done this way, but it's an ingredient that's added to our toolbox now, which is going to be important in concert with sometimes explicit objectives. It gives us a powerful new tool. This has actually led to a field called quality diversity, where we combine quality measures with diversity measures and try to do both at once in order to make a principled attempt to leverage what we know about both of those kinds of searches.

[0:44:33.4] SC: Super interesting stuff. Kenneth, I really appreciate you taking the time to speak with us about neuro-evolution and your research. Is there anything else that you'd like to leave us with?

[0:44:45.1] KS: Well, just to say that take a look at neuro-evolution, like it's actually becoming now more recognized in deep learning, that we have actually a lot of synergy with deep learning, because we're also doing neural networks. Both fields I think are realizing today that we have something to offer each other perhaps. Like neuro-evolution can evolve architectures and deep learning can apply really powerful learning algorithms to those new complicated architectures for just this one example.

Neuro-evolution can contribute to reinforcement learning in new ways, because of the way that fitness can be a different kind of driver of progress, than say the typical gradient-based approach. In the end, we get a possible really powerful synergy. I think it's worth looking at how these two things can possibly feed in to each other going forward.

[0:45:31.3] SC: Awesome. What's the best way for folks to learn more about what you're doing?

[0:45:34.5] KS: I'd point people to – I'm guessing you probably have some links associated with the interview.

[0:45:40.3] SC: Yup. We're going to include a link. I know you've got a page on the UCF site. Is that the best one?

[0:45:46.1] KS: Yeah. I'd point people to my homepage, my research group homepage, point there at UCF. Also, I can provide a link to Uber AI Labs, where we actually are hiring too. If people are just interested in jobs in general, that's another opportunity there. I'll also point to that.

[0:46:00.7] SC: Fantastic. Well, thanks so much Kenneth.

[0:46:03.1] KS: Yeah, thanks. It's been a pleasure. Thank you.

[END OF INTERVIEW]

[0:46:10.5] SC: All right everyone. That's our show for today. Thanks so much for listening and for your continued feedback and support. Thanks to your support, this podcast finished the year as a top 40 technology podcast on Apple Podcasts. My producers says that one of his goals this year is to crack the top 10. To do that, we need you to head over to your podcast app, rate the show, hopefully we've earned five stars, and leave us a glowing review.

More importantly, share the podcast with your friends, family, coworkers, the Starbucks barista, your Uber driver, everyone who might be interested. Every review, rating and share goes a long way, so thanks in advance.

For more information on Kenneth or any of the topics covered in this episode, head on over to twimlai.com/talk/94.

Of course, we would love to hear from you either via a comment on the show notes page or via Twitter to @samcharrington, or @twimlai, or @twimlai.

Thanks once again for listening, and catch you next time.

[END]

