Danny Hillis interview – Connection Machine Legacy. August 23, 2016

In this interview transcript, Tamiko Thiel interviews Danny Hillis, inventor of the Connection Machine and founder of Thinking Machines Corporation, on the work he and his fellow scientists did in the 1980s-90s and how that has influences contemporary parallel computing and artificial intelligence.

(Tamiko Thiel was responsible for product design of the CM-1/CM-2: http://tamikothiel.com/cm)

General interest links on Danny and the Connection Machines CM-1/CM-2, and the CM-5:

Danny's bio: http://longnow.org/people/board/danny0/

Connection Machine: https://en.wikipedia.org/wiki/Connection_Machine

The Evolution of Thinking Machines / Aspen Insitute talk https://youtu.be/LxqgGMmmj8A

Danny Hillis, Connection Machine Legacy: Sergey Brin/Google

https://www.youtube.com/edit?video_id=roCJ6tcku60

TT:

I was thinking, well it's 30 years since the machine was launched, a lot has changed in that time, at least in terms at least of what we can do with all the concepts for which the machine was invented. Hence my idea to start a project with sort of indefinite goals and timespan but just to at some level educate myself about what was happening in AI over the last 30 years.

DH:

OK, well to go back, one of the founding premises of Thinking Machines was that artificial intelligence would require parallel computers. And that's sort of an obvious statement now but if you recall, there were a lot of people who believed that there were proofs that parallel computers could never work. And the reason that I believed that they could, in spite of all the proofs, like you know, Amdahl's Law and things like that, was that I knew that the brain works, and it used very slow components, but a lot of them, so clearly a lot of slow components working in parallel were good enough enough for the computations that we were having trouble with, like pattern recognition.

Parallel computing: https://en.wikipedia.org/wiki/Parallel_computing

Amdahl's Law and where it broke down: http://www.essenceandartifact.com/2011/03/amdahls-law-is-wrong.html

And there was also a body of theory about how to make multi-level neural networks and train them to recognize patterns. But the computers at the time were so slow that you could only make tiny networks and it was very difficult to train them and so on. So it didn't really work that well, because of the scale - or I believed it was the scale - and so I thought that if you made a parallel computer and took advantage of semiconductor technology which was just kind of exploding at that time, that you would be able to make these pattern recognition kinds of things work and what you needed was something that would simulate connected networks of neurons which was why it was called a connection machine originally. And it was not just connected networks of neurons but semantic networks, linked things, … but the general quality was that information was represented in links and connections.

Neural networks: https://en.wikipedia.org/wiki/Artificial_neural_network

TT:

Rather than being stored in places or …

DH:

Right. And I had a programming paradigm for it, which is now what's called MapReduce, and so that programming paradigm was embodied in the hardware. So part of the hardware was a map hardware that broadcast out the functions to be mapped across everything, and then you put a different piece of data in every processor, you send out the function to be computed across it and then there was another network which did the reduce step to bring it all back together into something. And so the basic paradigm was MapReduce, and so that is indeed one of the biggest, or probably the main programing paradigm today for parallel machines.

Google's MapReduce: https://en.wikipedia.org/wiki/MapReduce

For mapping and reduction on the Connection Machine, see:
Hillis, W. Daniel. Chapter 2.2 Alpha Notation (mapping), p37, and 2.3 Beta Reduction, p41 and 2.7 "A Connection Machine is the direct hardware embodiment of the alpha (mapping) and beta (reduction) operators, p.47-48, The Connection Machine, MIT Press, 1985.

TT:

The other thing is that I know that MapReduce was the basic underlying part of the Google search engine which turned Google into, you know, started Google as the entity that we have today which is really dominating a lot of not only technical fields but our lives!

DH:

Well Sergey Brin was an early Connection Machine programmer, so it's no coincidence [laughs].

TT:

And I found his resume from that time online! http://infolab.stanford.edu/~sergey/resume.html

DH:

Oh that's funny! Yeah, he definitely used that and you know, made it apply to the Web, as Brewster Kahle was also trying to apply it to the web. But the Web didn't really quite exist when Brewster was doing it.

Brewster Kahle, CM lead engineer, WAIS, Internet Archive:
https://en.wikipedia.org/wiki/Brewster_Kahle

TT:

Was that in the '80s when Brewster was trying originally?

DH:

Yeah, Brewster was actually trying to do it sort of before the Web. He was trying to do it to the data that was out there on the Internet, but there was no common form for the data, the data was in files. So that's why he made WAIS which was a way of sort of looking at all those files and indexing them. So you could have a search engine before there was a Web. He was a little early in that, that really the search engine didn't really take off until there was a Web. Strangely enough, you know another piece of this was making big semantic networks and later I went back and redid that project after parallel computers came out, and that was what MetaWeb was about. Then Google bought that, and that's a big part of Google search now is that representation of knowledge, the knowledge graph, and in fact the CTO of MetaWeb [John Giannandrea] is now Head of Search [and AI research] at Google. So it all comes back full circle [laughs].

> WAIS Wide Area Information Server, Brewster Kahle/TMC: https://en.wikipedia.org/wiki/Wide_area_information_server
>
> Semantic networks: https://en.wikipedia.org/wiki/Semantic_network
>
> Metaweb, Danny's semantic networks company:
> https://en.wikipedia.org/wiki/Metaweb
> http://www.nytimes.com/2007/03/09/technology/09data.html
>
> MetaWeb's CTO John Giannandrea became head of AI research and now Search at Google
> http://www.thesempost.com/john-giannandrea-new-google-head-of-search-amit-singhal-leaving-google/

Danny Hillis, Connection Machine Legacy 2: Neural Nets and Semantic Networks
https://www.youtube.com/edit?o=U&video_id=IHTj4gjzFIA

DH:

So now, interestingly enough, Google is not only using that paradigm of MapReduce, but they're also using the semantic network, the parallel search algorithms on semantic networks and also they're a real pioneer on using the neural networks, for trained learning. That's maybe the most interesting story because that was something that people had been trying to do since the days of analog computers. It was well understood what the training algorithms were, and it was well understood that you needed multi-layer networks, and Marvin Minsky proved that you couldn't do much with one layer networks, that you would need more layers. But the training took much longer.

> Marvin Minsky: https://en.wikipedia.org/wiki/Marvin_Minsky

There are a bunch of people around, like Geoff Hinton, that kept working on it, thinking about how you would train them and so on. But the computational power just wasn't there. And it turns out it wasn't even there with the Connection Machine. They tried it with the Connection Machine, we made some progress and Thinking Machines used to have demos where it would recognize, you remember, like a hammer and a pair of plyers, and stuff like that. But it was pretty crude.

> Geoff Hinton: https://en.wikipedia.org/wiki/Geoffrey_Hinton

Then what happened was eventually, the big parallel computers became available through the cloud, and finally people could get - and through big piles of graphics processors, which are basically little Connection Machines …

> The Cloud: http://gizmodo.com/what-is-the-cloud-and-where-is-it-1682276210
>
> GPUs (Graphic Processing Units): https://en.wikipedia.org/wiki/Graphics_processing_unit

TT:

Right - Nvidia's GPUs

DH:

… like the Nvidia things and so on. So that it became possible to really run these much much bigger - literally thousands of times faster than we could do on the biggest Connection Machines. It turns out when they ran that on the data, and the data was much more available then, the algorithms *worked* and so all the things that we used to talk about, like recognizing faces, and telling cats from dogs and so on, are basically are now known how to do using neural networks. So the pattern recognition piece of intelligence, I think, is kind of a solved problem if you have a big enough computer. I don't think that's the only thing there is to intelligence though.

> Pattern recognition: https://en.wikipedia.org/wiki/Pattern_recognition

TT: So what are the other things?

DH:

For instance on the symbolic side there's also big progress, looking through big semantic nets, the knowledge networks, so when you do a Google search now, it actually does two searches. It still does the old keyword search, but it now uses the semantics of it, the actual meaning of things.

TT:

But the semantics network is trained through neural networks, isn't it?

DH:

Has changed what?

TT:

The semantic networks, the semantic networks had to be trained through neural networks in order to ...

DH:

Well, no, that's actually not true, those are really two different things and two different searches occurred. So one of them is happening at the level of symbolic processing. The other is much more ... it's almost pseudo analog processing.

You know, it is clear that there's lots of different kinds of intelligence but, roughly speaking there's one kind of intelligence that corresponds to very fast kind of recognition and intuition, where you have lots of weak data, you integrate it, and you come to a conclusion. So that's really what neural networks do very well: combine a lot of weak data together.

And then there's a completely different kind which is much more reasoning which is you have chains of facts and symbols and it's much more what we do when we do explicit conscious reasoning about things. And the semantic networks lend themselves more to that, so they're really symbolic processing. So both of those things, it turns out, use parallel computers very well. And those have made some real progress lately, although the neural networks are the ones that have obviously made the biggest jump recently.

Danny Hillis, Connection Machine Legacy 3: More on Semantic Networks
https://www.youtube.com/edit?o=U&video_id=RKnCrj6riVE

TT:
The semantic networks sounds like knowledge systems that people like Doug Lenat were working on. Is that true?

> Doug Lenat, Expert Systems - Cyc:
> https://www.wired.com/2016/03/doug-lenat-artificial-intelligence-common-sense-engine/

DH:
Well they were, except that they were much more general ... those expert systems had very specific knowledge of a very limited domain. Things like the knowledge graph literally knows about hundreds of billions of entities. I mean they know about you, they know about the building that you're sitting in, they know about the computers that we're using, they know about the language, the words that we're using as entities. So they literally know hundreds of billions of facts, and are able to make inferences about those facts, connections between those facts and so on.

> Google Knowledge Graph:
> https://www.google.com/intl/es419/insidesearch/features/search/knowledge.html

Now, in principle it may be that you can encode that kind of information in neural networks, and I think that might be one of the next things that happens in neural networks as people start making a bridge between them. But today that's not mostly what people do with neural networks. What people do mostly with neural networks is more kind of recognition tasks, of taking patterns of sensory inputs and putting them into categories.

TT:
Marvin once said if we connected a machine or a robot to a TV it would be as smart as a human, which I took to also be a derogatory statement about what he thought of the intelligence of most humans! But ... do you know that line of his and it did it mean ...

DH:
It sounds very much like something Marvin would have said! I don't specifically remember him saying it, but that would have been the sense in which he would have meant it.

TT:
I took it to mean also that most people get their knowledge about the world through something like television. Some of us even read books, but your generic person gets it through television, or of course also through contact with the world and other human beings. But any sort of larger knowledge source was for a lot of people the TV. And so I was wondering if he was talking about these sorts of knowledge bases, essentially.

DH:
Well ... I don't know, Marvin was very much into common sense knowledge and how did you represent them. A lot of Society of Mind was how do you do things like tell the difference between a dog and a cat and so on. I think the stuff that was apparently so easy for us like that turned out to be much harder than the stuff that was apparently very hard for us, like playing chess.

> Society of Mind, full online text and illustrations: http://aurellem.org/society-of-mind/

TT:
But he emphasizes that - I've been rereading the Society of Mind - he emphasizes over and over that he thought that what seemed to be simple for us is the most difficult for machines, and that's where he throught the real work would lie, not in expert systems being able to deduce things.

DH:
That's right, and I think he has been proved right about that. You know, most of AI is in the difference between a toaster and a two-year-old, not between a two-year-old and an expert, right? It's exciting now because we have in that specific area of recognition made some huge strides and it's gotten people very interested in AI again as a field and it's got lots of resources for the area. You know, there's kind of a cycle that happens with these things, and so there's some progress and then people think everything is solved, and then ... everybody starts wringing their hands about, you know, what happens when our AI overlords become so much smarter than we are and then, we discover, actually there's a lot more that we don't understand about intelligence than what we just learned. I'm old enough to have seen three cycles of that now.

Danny Hillis, Connection Machine Legacy 4: Emergent Behavior in Human and AI systems

https://www.youtube.com/edit?o=U&video_id=HgX1yNTEq7k

TT:

Let's back up a bit and say what does interest you on AI, what do you think are the exciting things for you personally - maybe more than for the field in general?

DH:

Well, I guess for me personally, I think the most interesting and most legitimately frightening, but also most legitimately impactfull thing is when machines learn how to improve themselves. And not just learn within a structure set up by humans but to actually build new structures and learn how to learn better and so on. So as soon as the machines can start bootstrapping their intelligence, then that potentially becomes extremely interesting because it's no longer governed by the speed of human understanding and development. Because right now it kind of is.

> Machine learning: https://en.wikipedia.org/wiki/Machine_learning
>
> Guardian journalist tries out machine learning:
> https://www.theguardian.com/technology/2016/jun/28/google-says-machine-learning-is-the-future-so-i-tried-it-myself
>
> Computers that can teach themselves: DARPA project 2013 - 2017:
> Probabilistic Programming for Advanced Machine Learning

https://www.wired.com/2013/03/darpa-machine-learning-2/
https://galois.com/project/probabilistic-programming-for-advanced-machine-learning/

Oh, a window just popped up right in front of me. Sorry about that, I have to deal with it. Okay. For some reason my machine decided to get a picture of this right now. What can I say. So for instance the window that just popped up is for instance a CAD program.I''m still telling it shapes, and putting the shapes together and so on, but very soon I'm sure that I'll just tell it goals, like I want these two things be held apart with this force. Already there are some CAD programs that do that.

So one of the things happening is even the design of our society, the design of our machines themselves is much less connected to human understanding, I mean it's already so, nobody really understands in detail how the Internet works, for instance. They understand parts of it, they understand globally the big picture, but it does things that nobody would have predicted. And so our relationship to it becomes more like a relationship with a natural object. But nobody's really in charge of its design.

Tinker With a Neural Network (Google Tensor Flow):
http://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=4,2&seed=0.17225&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false

TT:

One thing that Marvin talks about in Society of Mind is that learning is one part of it, but learning still needs goals, and the goals will be set by humans. But what you just said about the Internet developing structures that people don't quite understand seems to imply that even if we build things with certain goals, then at some point … structures can start evolving that are not focused or directed by those goals.

DH:

Yup! That's true and we already have that problem. We already have that problem for example with organizations. So when you start a company you may have certain goals but after a while the company has its own goals, or the NGO has its own goals, or the country, yeah, the political organization. And so things have goals, they develop goals of their own which are not necessarily the goals of the designer. They are emergent goals and in fact in many ways, end up having goals that are pretty contrary to the goals of individual humans! They naturally want to perpetuate themselves and gain more power and so on. So I think some of our concepts like democracy are getting strongly corrupted by the fact that we've made these entities which have in some sense super human capacities, and another sense some human capacities, but they are already very powerful actors in our lives. In the same way that we try to control them with only partial success I think probably the same thing will happen with our machines. I think that's a very good analogy.

TT:

That seems to imply that we should be studying these large organizations and how they function and how we can direct them in order to be able to - I mean that will happen more and more.

DH:

Yeah that might be a prediction and maybe … of course they're different in the sense that we have different control models to use, because they're made of organizations of people, they have to be controlled with the tools that influence people, like leadership and persuasion and things like incentives. Those are ways that we try to influence those organizations. But we'll probably have to use some of those tools, but we'll have to use some different tools to influence the machines. So I certainly believe that it's going to create problems for us, but I also believe that it's ... it's probably going to be the thing that solves a lot of our problems too, if that make sense.

TT:

Yeah or changes the types of problems that we are confronted with.

DH:

Yeah. It changes … it's fundamentally going to change who we are, just like civilization has changed who we are. And some ways for the better, some ways for the worse. I do believe we're going through a kind of revolution that I sort of make the analogy to the Enlightenment. I think the Enlightenment really created the idea of the individual as, you know, the ultimate creation, act of creation and the individuals, the autonomous agent that could perceive and build and understand everything. And I think that what we're going through now is something I call the Entanglement.

> Danny Hillis: “The Enlightenment is Dead, Long Live the Entanglement”
> http://jods.mitpress.mit.edu/pub/enlightenment-to-entanglement

TT:

I've read your post on the new Media Lab blog.

DH:
Yeah, I think that it's all about connections, and as we become much more connected … it's this connected emergent thing that makes for instance these moral decisions ... it's the things that have goals or combinations of people and machines ... and are tied together by much tighter communication and much more connectivity. And so it brings us into a different kind of world that is potentially much more ... intelligent ... in the sense that it's computationally much more robust and it can fix other problems. It's not necessarily .. more intelligent in all dimensions though.

TT:

Well, and intelligence is ... is potentially very different from …

DH:

Wisdom???

TT:

Good, or positive, or human, or good for humans or ...

DH:

Yeah! Yeah that's right, and in fact I think ... I guess it's only optimism that makes me believe that there's some positive progress in this. I don't see any guarantee ... it becomes this kind of existential question, but I don't think there's any guarantee that something smarter would be better! And it might be also: in whose eyes better? In Brave New World when you took the people from the past, they thought the future was much worse. But the people from the future looked back at the past and thought it was much worse. And certainly I wouldn't want to be transported back and have to live at any time in the past.

TT:

Yeah me either, I mean especially as a woman frankly.

DH:

Well, that's a perfect example I do believe in the the Martin Luther King comment that "the arc of history is long but it bends towards justice." And I do believe in a sort of slow moral progress ... BUT it's … I think what the question is, is *why* does it bend towards justice? And my hope is that truth has a certain ... that there is an underlying... truth to things, and that things that are resonant with the truth, survive for longer than things that are counter to the truth.

Danny Hillis, Connection Machine Legacy 5: Cloud, GPUs, People
https://www.youtube.com/edit?o=U&video_id=Jwc6A3Posro

TT:

You had mentioned I think when I was visiting you last, that the Connection Machine you saw as sort of a dead end. Were you talking about hardware? Because what you have just told me is that all the concepts that - I mean I know that you were, you wanted ...

DH:

Well I don't think it was a dead end, I think it was ahead of its time, and the technology wasn't quite there to do what we wanted. But it definitely showed that something like that was possible, and it got people thinking in a way that as the technology came along, that you could build, you know, you could do that economically on a large enough scale that people sort of knew how to program it, knew how to use it, and so on.

I remember when I wrote the first Scientific American article on the Connection Machine, I had an hypothesis, I said look, once we have this, there's no need to have the power of the computer in your home so much. The real power will be out in some centralized location. I didn't call it the Cloud but that's where you need computations, because computations want to be near each other. They don't need to be near you once you have the bandwidth for that. And I said so you can imagine you have big computers that you use, something across the nation, and then the Scientific American people made me - they were like, "that's just too implausible. We'll let you say, like, 'in the same city.'" So I was like, "Okay, well fine." So you know, one city, someplace in the center of the city, but they actually - that was the negotiation of them pushing it back because it just sounded too crazy.

> Hillis, W. Daniel. "The Connection Machine," Scientific American, 1987
> https://www.scientificamerican.com/article/the-connection-machine/

TT:

Wow. Well, can we say that the CM-1 as a SIMD machine, that the hardware and software experience there flowed into the GPU processors. And the CM-5 ...

SIMD (Single Instruction Multiple Data): https://en.wikipedia.org/wiki/SIMD

DH:

Yeah. A GPU chip or even a microprocessor that has a GPU unit or a vector unit on it, you know, they have SIMD instructions on them, so those are fundamentally little Connection Machines and I think probably people like Nvidia even acknowledge that there're little Connection Machines on the chip, and they can put several - they actually don't have all the things the Connection Machine had, because they don't have the connection part of it so much, but they have the map part more than they have the reduce part. Or they have the connection part, the communicate part. They're able to do reduction. ...

So that's one sense in which it exists, and then the Cloud is much more like the CM-5 it's racks and racks of micro processors that are connected by a fast network. And so that was very much the architecture that we evolved toward, and they're programmed very much like the CM-5 was programmed. Although they actually don't have special hardware for reduce network anymore they just do that with the general network. And they don't have special hardware for the map like the graphics processors do. So it turns out for the neural networks actually the more efficient way to do it is on a bunch of the graphics processors because they do have that special hardware. So both things definitely take elements from the way that we did things on the Connection Machine. It's interesting though, that in many ways the software is more primitive than then ... we really had it so that you can take your Fortran program and compile it to run on 10,000 processors and that pretty much doesn't exist now.

TT:

Why *not*???

DH:

The technology kind of lost in the whole meltdown of the super computer industry ... there are DARPA programs to deal with it again. But it's funny seeing a technology get lost in your lifetime. It will get rebuilt eventually. But people were able to do enough with the simpler programming paradigms, the simple kind of things like MapReduce, which was definitely one of the things we used but wasn't the only thing we used. But that's turned out to be powerful enough to do an awful lot.

TT:

So it sounds like even though Thinking Machines went belly up after little over 10 years that in some sense it was worthwhile, both the hardware and the software and the ideas have gone on.

DH:

Well also the people have gone on, I mean it's amazing … some of them have gone on to win Nobel Prizes, or start giant research institutes, or found companies or ... So what's interesting is actually what would have been the best investment portfolio would've just been to invest in everybody in that building - whatever they did! And so, you know, fantastic people if you remember, like Eric Lander was just starting to play with biology using the [machine] ... and is now the head of one of the biggest biology institutes. Sydney Brenner... only insiders knew who Sydney Brenner was, he certainly hadn't won the Nobel Prize at that point.

> Eric Lander
> http://link.springer.com/article/10.1007/BF00128166
> http://www.yourgenome.org/stories/giants-in-genomics-eric-lander
>
> Sydney Brenner
> https://www.nobelprize.org/nobel_prizes/medicine/laureates/2002/brenner-bio.html

So it was things like Brewster's search engine and his archiving the Internet just seemed like crazy ideas, nobody understood what they were. A lot of people went on to do kind of amazing things... so maybe that's how it has it's biggest impact, was a set of people - it's like the Manhattan Project in that sense, a set of people got together, and inspired each other to go off and do great things.

> Wayback Machine / Internet Archive: https://archive.org/web/

And yeah we did do a bunch of stuff and we did - I mean, the project that Dick Feynman was working, on which was quantum computing, and it was so absurd that only Dick Feynman was working on it - that's become a whole field! And I'm quite sure that Thinking Machines was the first company that ever worked on that!

> Richard Feynman:
> Connection Machine: http://longnow.org/essays/richard-feynman-connection-machine/
> Quantum Computing: https://people.eecs.berkeley.edu/~christos/classics/Feynman.pdf

TT:

Just because they had this nutty guy called Richard Feynman who had this nutty idea that he was working on in his spare time.

DH:

Yeah, yeah, yeah. So well I think it was ... I think we were very lucky to have been at that time in that place, and it was certainly an extraordinary thing. If I had known even a quarter as much about this business and how the world works as I do now, it would have been able to save it for long enough to, you know, for the web come along and parallel computing to come along, but you know, I was a bad business person, made a lot of stupid mistakes and you know, I can now go back and see all kinds of things that I did wrong that I would never do today, but it wasn't - it wasn't what I knew about, it wasn't what I was paying attention to then.

TT:

Yeah and it did go belly up right when the web started out didn't it, that was that was a bit of a ...

DH:

And remember, we were working on a web server and nobody - we couldn't raise any money on it because nobody knew what the web was, right? Right. Sigh. But in fact, you know, a lot of those people went to Sun, and turned it - helped turn it from a workstation company into a web company.

Sun Microsystems Purchases Thinking Machines' GlobalWorks
http://www.eetimes.com/document.asp?doc_id=1209360

The Path to Posthumanity, 2006
https://books.google.de/books?id=_YkfoiKC4PcC&pg=PA126&lpg=PA126&dq=sun+microsystems+and+thinking+machines&source=bl&ots=PlsgmG6GWa&sig=-uTr5Rof9YtvDvH-cWyamT3t73c&hl=en&sa=X&redir_esc=y#v=onepage&q=sun%20microsystems%20and%20thinking%20machines&f=false