

A UBIQUITY INTERVIEW WITH DAVID HANSON

[As CEO of Hanson Robotics, Inc, David Hanson creates robot faces that have been dubbed "among the most advanced in the world" by the BBC, and inspired Science to label Hanson "head of his class" in social robotics. After receiving a BFA from the Rhode Island School of Design, and dabbling in AI programming at Brown, David Hanson worked at Walt Disney Imagineering, leading development of an autonomous walking robot and electro-active polymer (EAP) actuators. Later Hanson went on to work toward a PhD at the University of TX at Dallas, developing social robots affect naturalistic conversations with face tracking AI, speech recognition, and realistic expressions that use Hanson's patent-pending polymer materials.]

UBIQUITY: Let's start by discussing some of the work you and Mihai Nadin have been doing. [Readers who missed our interview with [Professor Mihai Nadin](#) may want to read it in the Ubiquity archives.] What's the focus of your research?

HANSON: We're looking at how robots can be used to increase in the anticipation skills of the elderly. The idea is to create small characters that engage people in a very personal way, since people respond naturally to conversational characters -- be it ones they encounter in a work of literature, in a screen animation, interactively in the form of a person.

UBIQUITY: Give us a little lesson. Your specialty now is social robots. What other kinds of robotics are there?

HANSON: I built a small walking robot back in college at RISD taking classes at Brown. I was taking a couple of computer science special topics classes, and I was looking at using robotics at the AI to control themed environments and for things such as looking at how to induce playful states in people's minds. So my focus is on the concept of what I call robotic party architectures, where the idea is to create a space that's really playful and surprising and then invite people to it under the premise that it's a party. But then it's more like a sort of psychoactive themed environment that takes you on a voyage of sorts. And I regard that kind of voyage as a recurring effect of art. Good art often seems to transport you to someplace new or to a new state of mind, expose you to new ideas that arose during the course of the making of the art, or to convey ideas or they inspire new ideas in the viewer. So the viewer winds up having these playful changes of mind.

UBIQUITY: So what would your definition of social robotics be and how does it differ from other kinds of robotics?

HANSON: Social robotics is comprised of robots meant to engage people socially. The idea is that you have a set of normal mechanisms for engaging with other people socially. We've evolved to that way. And if you can create an artificial entity that utilizes

those neuro-mechanisms, then it becomes very natural for people to interface with a social robot. Those neural mechanisms function to extend our individual intelligence into a social intelligence; in other words, we're smarter when we organize them into social groups, like families and friendships and corporations and government and schools. These are all social institutions that are built on our natural tendency to socialize by these very primal neural mechanisms. The idea is that by tapping into these tendencies, we are smarter when we interface with our machines.

UBIQUITY: How and why does that happen?

HANSON: Our machines become smarter when they interface with us because in order for them to use our natural social interfacing skills, they have to become smarter. They act socially smarter. They begin to mirror our neural mechanisms, not necessarily exactly, but in some kind of functional way. We have wet neurons that are making the robot smile at the appropriate moment, and if it smiles at the appropriate moment then you're starting to approximate social mechanisms.

UBIQUITY: You've done some amazing work, and we need to send people to your Web site. What would be the thing that you would start showing people from your Web site? What should they look at first?

HANSON: I would encourage them to look first at the [Hanson Robotics site](#) to look at the Einstein video to see how good the hardware is, how lightweight and low-power it is. And then after . the Einstein robot the [video of the Philip K. Dick robot](#), which shows the conversational capabilities of our robots, the ability of the robots to capture and hold people's attention in a conversation. There are so many uses for social robots. One is entertainment. I mean, we're fascinated, we gravitate towards the human face and to human-like characters. The Philip K. Dick robot is a machine that literally holds a conversation with you, looks you in the eye and has an open-ended conversation, where you can talk about just about anything.

UBIQUITY: How do you do that? What's the trick?

HANSON: Well, there are a bunch of tricks. One of the weird things about a robot is that you can't really reduce it to one thing, because it's a lot of things coming together. It's an integrated system, so you have to have speech recognition, some natural language processing, some speech synthesis. You have to have computer vision that can see human faces and you have to have a sophisticated enough motion control system that you can direct the robot's eyes to the coordinate position where you've detected a face. And you have to remember where the last face was that you were looking at, you have to be able to look back where the face was. Then you start to hit on the rudiments of a world model, that three-dimensional world model that's populated

with some kind of representation of objects, people, and places.

UBIQUITY: How do you approach the problem?

HANSON: We're starting to sketch the requirements for the base conversational systems. Earlier we did this with less deep conversation. We did it with a robot called Hertz and a robot named Eva, and then finally we got to the Philip K. Dick robot. But the main difference is the depth of his conversational database, and to achieve that we did two things. First, we scanned a little over 10,000 pages of the writing of Philip K. Dick, and then developed a statistical search database similar to what you might have when you go to a search engine like Google and you type in a query. But ours is tied into natural language processes, so it has a special way of interpreting your question by parsing it into a search going to the database, semantically searching and then assembling the search results into a sentence.

UBIQUITY: Did you use off-the-shelf software for that or did you have to develop it all yourselves?

HANSON: We're using mostly off-the-shelf software. The speech recognition was provided by Multi Modal. The search software was provided by a company out in Colorado, using a technique called latent semantic analysis. The roles engine that we were using is Jess, which has an open source version, but it's based on Java. The natural language processing system was developed by

Andrew Olney, a Ph.D. student at the University of Memphis, who actually swept through our previous code and essentially rewrote all the language stuff from the ground up.

UBIQUITY: Where do you see your research going?

HANSON: I'm interested in making these robots easily custom-designed and mass producible -- in other words, easily designed using low-cost hardware, so that very inexpensive facial expressions can go with inexpensive walking robot bodies, as well as easily customized software. Therefore, we will be improving the software, improving the quality and rate of the speech recognition. The ability to design a custom personality and animation for the robots and to tweak and tune those things needs to get better. I see these as practical tools for bringing social robots into our lives, be they human-like or cartoon-like. These tools will be useful for artificial intelligence development. In an essay a couple of years ago AI pioneer Marvin Minsky lamented the fact that the graduate students and the AI lab at MIT had spent most of their time soldering instead of developing artificial intelligence.

UBIQUITY: Did he propose anything to fix that problem?

HANSON: His solution was to turn to simulation, but the problem with simulation as far as characters are concerned -- or as far as social agents are

concerned -- is that you lack that sense of "presence," that sense of immediacy that you gain when you have a 3D robot. Also, a virtual world doesn't have all the physics of a real world, so you're not simulating the noisiness of the environment. Of course, we say "noisiness," but it's really information -- just not information that we understand necessarily. If we lay out these tools for software and for hardware, rapid design and low cost deployment, then we get rid of those obstacles to AI development. If we can provide these software tools that are easily extensible, modular, and replaceable, but which provide a foundation for AI character development, then it means that the research community can focus on other problems.

UBIQUITY: In what way?

HANSON: The research community can turn into an extended development network so that the parts and pieces will play well together. The human-like interfaces and the interactive software systems can operate with an animated agent. This way, researchers can focus on hard problems in a subdomain, like speech recognition or general intelligence, but investigate how these components operate in the real world, in face-to-face conversations with people, while walking around the world. Solving the hard problems of AI may not always require a virtual agent or a physical robot body, but I see the physical robot bodies being very useful in research, and certainly useful in the marketplace. For example, I see animated robotic

characters -- think Disney characters or Universal Studios characters -- that live in the house and actually walk around. They look up at you, they lock eyes with you, they converse with you, they can teach your kids, they can provide cognitive stimulation for the elderly, they can download wirelessly the latest news and gossip. They can be an extension of the world of entertainment, and can also be a natural language search tool, so that if you have a question you ask the robot. The robot sends the question off to the search API, and turn it into a natural language interface for the robot. (That hasn't been done yet, but it can be done.) And then the robot returns the answer to your query in natural language.

UBIQUITY: How good is this natural language?

HANSON: Well, that's one area that requires more development. For precise uses of the robot -- for example, if you wanted to know the capital of the country -- you would need the speech recognition to be spot on. Unfortunately, recognition with a large vocabulary is about 50 to 80% accurate, and that's the good news -- but remember that if it's misunderstanding 20% of the words in a sentence, that's the bad news, because then the whole meaning of the sentence can be lost if one in five words are misunderstood. So speech recognition is one stumbling block right now, and another stumbling block is coming up with a good interaction design that solves some of the problems with speech recognition. Even with the speech recognition of

human beings, we sometimes don't understand each other. We have to have little dialogue techniques for asking what the other one was meaning or what the other one said. If you don't understand something, then you pipe up and say something, so you can begin that disambiguation routine.

UBIQUITY: On the question of interaction, can the robots have social interaction with each other?

HANSON: Yes, absolutely. The artist Ken Feingold created a work with two chatterboxes side by side in a box, for the Whitney Biennial in 2002. Two talking heads just spit responses back and forth to one another. Of course, chatterboxes are very shallow natural language interaction devices, and all their responses are pre-programmed, and they don't do any semantic evaluations. It's usually a one-for-one type response, so it's predictable and shallow; however, they are still very interesting. The conversations you can have with a chatterbox can be very stimulating. Feingold built two identical robots and then fed the text from one chatterbox into the other chatterbox, so they were in this infinite loop of chatting with one another. And the results are interesting. But as we build social interactive robots with deeper minds, deeper natural language capabilities, and more open-ended capabilities, then the interaction between the robots or artificial minds will get more rich and interesting. I believe that that human intelligence is substantially social intelligence: we work well socially and we solve problems together that we can't solve alone,

and I believe that robots and artificial minds will work in concert in a similar way. If we're emulating our own social intelligence, then the machines will get smarter in kind. They'll be smarter as a community of robots.

UBIQUITY: How many people and how many people hours do you think were devoted to -- just to pick one project -- let's say the Philip K. Dick project? What went into that?

HANSON: We put in probably 200-300 hours on software development in Dallas, plus the foundation of the software development took thousands of hours, and the work on the robots prior to the Philip K. Dick robot. So you've got all that heritage development work. But then from the point where we hit Go on the PKD robot, when I started sculpting it, I personally put in probably about 800 hours. We had a CAD designer working on it who put in maybe 120 hours, and he was taking all the parts and pieces that were made by hand in previous years and turning them into CAD models that could be produced on a prototyping machine. We had about four other people. I think each one of them probably put in 200-300 hours. Andrew Olney, as I understand it, put in a couple of hundred hours on it.

UBIQUITY: Let's turn from software to hardware -- specifically, skinware. Is there something called skinware?

HANSON: Sure, sure. Well, that would be part of the hardware, but yes, absolutely.

UBIQUITY: It's really impressive how lifelike you've been able to make the skin of the robots.

HANSON: Oh, thank you, yes. The facial expression robots like the ones that you might see in Japan or in animatronics for theme parks and motion pictures used the best material available, which was rubber material, an elastic polymer really. Rubber refers to like latex rubber, so usually they say elastomer, or elastic polymer, and the physics of solid elastomers doesn't behave like the physics of facial soft tissue. Facial soft tissue is a cellular elastic material mostly filled with liquid probably, 85% water. It winds up taking considerably more energy to achieve compressive displacement in the material than to elongate, to stretch it. So the key in my mind was then to achieve some kind of cellular grid or matrix in the material, but the conventional sponge materials have stress concentration points, which just means that where the cell walls meet, they get thicker, the material gets thicker and it stretches less.

UBIQUITY: What's the consequence of that?

HANSON: The most important is that the material wrinkles, because wherever you have a wrinkle on the face, let's say that crease around the corner of a mouth when you smile, that is an area where the material has to compress. In order for it to fold, it

has to compress right there along that line, and conventional elastomers just won't do it. It's very difficult to get the kind of wrinkling and creases that are natural; even in child faces you have those creases, you know? It's not just older faces that have those hallmarks of human expressions, so by making the material foam, we're able to get those good expressions and we're able to achieve the expressions with substantially lower energy.

UBIQUITY: Look back on the history of artificial intelligence and social robotics and help us see it as a unified history. You remember Eliza, right? Start from Eliza, and tell us what's happened since then.

HANSON: Well, that was the first chatterbox and it functioned very, very well in a simple way. But, as I said, the intelligence that you can embody in a chatterbox is going to be shallow. There is no world model, no semantic analysis of language. There was a standing bias against anthropomorphisms that was kind of inherited from the scientific wariness of the anthropomorphic bias, where you wanted to not project the human mind on animals or on anything in nature. So the idea was when you make robots, you don't want to project your human biases on these things, you want the intelligence to sort of exist independently of human origins. This model was reflected in the movie 2001, with the Hal character. So Hal doesn't have a face; it has only a voice and just this one eye. It's extremely creepy, to be honest. But the development of the Hal character in the novel and the movie was based on consultation

with leaders in the field of AI. And the kind of robots that you might see in movies like Westworld or in Blade Runner was not taken seriously in the world of robotics.

UBIQUITY: When did this approach change?

HANSON: The idea started changing in the '80s and '90s when you started having this concept of a mode of computing or affective computing. The idea was that computers could benefit from having emotions in two ways, first by being easier to relate to by humans, because if a computer seems emotional then it's not as cold: it makes you feel better, it makes you feel like it likes you -- and that in turn can help you like it. And then the second way is the way that Antonio Damasio describes the world of emotions in human cognitive processes in his book called "Descartes' Error." Damasio's a neuroscientist who has investigated the role of feelings and emotions in humans -- in particular, humans who have had brain damage to areas related to feelings and emotions and whose performance on IQ tests and all kinds of other tests is fine but they can't perform well in the world: they screw up all their relationships, their jobs, they can't calculate things like the odds of success; in long-term gambling scenarios in human subject tests and in the tests where they have to make decisions about long-term gains versus short-term rewards, they always choose short-term rewards. By making computers affective, we allow humans to relate to the computers better

and the computers to perform in the world better. That's the premise.

UBIQUITY: Where did we go from that premise?

HANSON: With that model and with the work of Cynthia Breazeal at MIT, you had a birth of social robotics -- this idea that you can make the characters have a face that looks very humanlike in a simplistic way, with frowns, smiles and facial expressions that we can recognize. And then you go to systems where you look at the robot, the robot looks at you, you look away, the robot looks where you're looking, so that you are simulating this feeling of the robot's perception. People feel that the robot is intelligent. You can use that shared attention system then to begin to teach the robot, so the idea is that you have intelligence that emerges and is trained the way that a human infant would be trained. That kind of idea of an emerging intelligent system kind of sparked a fever of research in the mid '90s and Cynthia Breazeal has continued to have good results. Her latest robot, Leonardo, is proving capable of some pretty neat stuff -- learning names of objects spontaneously, learning simple tasks, physical tasks like flipping on switches and moving objects around in a three-dimensional space. But where that is going to lead is very encouraging. I mean, if you think of human beings, our intelligence is partially sort of prewired with our physiology, but then substantially programmed by interactions with our parents and teachers. So if you're trying to make an entity that is as smart as a human being, it would

make sense that it would emerge from ground up, but I think that we can also start to go from top down.

UBIQUITY: Expand on that point.

HANSON: I mean, we have all these intelligent systems that can do all kinds of things we consider intelligent—they perceive faces, can see facial expressions, can understand natural language to some extent, that can do sophisticated searches, that can perform as expert systems during medical diagnostics, et cetera. And if we patch all those things together, we can wind up with something that is very smart, albeit not as smart as a human being. This is the top down approach. In the short term such a system may simulate the intelligence of a human well enough to perform tasks and teaching and entertainment to the point where it can be useful in our daily lives. It is not as smart as a human, but acts smart. Yet the top-down folks hope that in time, emulating the highest apex of intelligent human thought processes, will result in truly strong AI.

But this contrasts starkly with the bottom-up approach which is to build robots that can't understand speech or do anything so smart in the short term, but are bug-like or babylike, under the hopes that they will evolve the high-level intelligence over time-- sort of recapitulating phylogeny of mind. I think that the two approaches will converge at some point in the future, so that you wind up having these entire levels of functionality by cobbling

together these chunks of solutions, but then that continue to evolve and learn over time because of the brilliant contributions of researchers like Cynthia.

[END]

Ubiquity -- Volume 7, Issue 18 (May 9, 2006 - May 16, 2006)

<<http://www.acm.org/ubiquity/>>.