

# Interview: Dr. Ben Goertzel on Artificial General Intelligence, Transhumanism and Open Source (Part 2/2)

June 13 2011, by Stuart Mason Dambrot

---



Dr. Ben Goertzel (right), Stuart Mason Dambrot (left). Photo courtesy: Neural Imprints (<http://www.neuralimprints.com/>)

(PhysOrg.com) -- Dr. Ben Goertzel is Chairman of [Humanity+](#); CEO of AI software company [Novamente LLC](#) and bioinformatics company [Biomind LLC](#); leader of the open-source [OpenCog](#) Artificial General Intelligence (AGI) software project; Chief Technology Officer of biopharma firm [Genescient Corp.](#); Director of Engineering of digital media firm [Vzillion Inc.](#); Advisor to the [Singularity University](#) and [Singularity Institute](#); Research Professor in the Fujian Key Lab for Brain-Like Intelligent Systems at [Xiamen University](#), China; and general Chair of the [Artificial General Intelligence Conference Series](#). His research work encompasses artificial general intelligence, natural language processing, cognitive science, data mining, machine learning,

computational finance, bioinformatics, virtual worlds and gaming and other areas, Dr. Goertzel has published a dozen scientific books, 100+ technical papers, and numerous journalistic articles, and the futurist treatise *A Cosmist Manifesto*. Before entering the software industry he served as a university faculty in several departments of mathematics, computer science and cognitive science, in the US, Australia and New Zealand.

Dr. Goertzel spoke with Critical Thought's Stuart Mason Dambrot following his talk at the recent 2011 [Transhumanism Meets Design Conference](#) in New York City. His presentation, **Designing Minds and Worlds**, asked and answered the key questions, *How can we design a world (virtual or physical) so that it supports ongoing learning and growth and ethical behavior? How can we design a mind so that it takes advantage of the affordances its world offers?* These are fundamental issues that bridge AI, robotics, cyborgics, [virtual world](#) and game design, sociology and psychology and other areas. His talk addressed them from a cognitive systems theory perspective and discussed how they're concretely being confronted in his current work applying the OpenCog Artificial General Intelligence system to control game characters in virtual worlds.

---

This is the second part of a two-part article. The first part is available at <http://www.physorg.com/news/2011-06-dr-ben-goertzel-artificial-intelligence.html>

---

**SM Dambrot:** What's your take on the [Blue Brain Project](#)? They've apparently emulated a cat's neocortical structure and announced that their goal is to emulate a human neocortex within, at this point, roughly eight years.

**Dr. Goertzel:** This is a long and complex story regarding a number of fascinating simulations done on IBM supercomputers. If you look at

what [Henry Markram](#) did in simulating a cortical column, in the Blue Brain project, that was very interesting from a number of standpoints -- yet in some ways it didn't do everything some people think it did. In simulating that column, Markham had to dig deeply into the equations of the flow of charge along a single neuron – and he actually published some really cool papers in [Biological Cybernetics](#) about adjusting those equations based on the measurements he and his team made. On the other hand, when you look at what the actual simulation he ran was, you can see that they did not actually simulate the precise input/output behavior of the cortical column.

What you'd like to see ideally is a simulation where if you feed some input into the column and get some output from the column, you see exact agreement with what you'd get from a real cortical column. They didn't do that; what they did do was create a simulated column that statistically had the same input/output properties as a real column. That's worthwhile and interesting, but it's not uploading a cortical column. Since we don't know the information coding of the column's inputs and outputs, we don't really know if we've gotten everything that's there. Imagine that you simulated the input/output properties of me as a language user in this way: from the statistical standpoint of acoustic analysis it would look like it had the same input/output properties as I do, yet it's missing the information.

Now, the cat brain that you mention was actually [Dharmendra Modha's](#) work. It was a totally different project based on IBM hardware that was the next generation from what Markham used. They simulated a neural network similar in size and connection complexity to a cat's brain. However, the pattern of connections was random – not derived from study of the cat brain and it didn't go down to the level of neurotransmitter concentrations either. It was a wonderful hardware demonstration of building a formalized neural network of that huge size, but it didn't have the same dynamics or structures as a cat brain because

we don't know what those are.

As it happens, Modha's team at IBM has done some other work aimed at understanding those structures, and published quite an interesting paper on the structure of the monkey brain in which they curated thousands of neuroscience papers and charted which regions of the monkey brain connected to other regions, trying to parse the connection structure just on a region-to-region level. There are hundreds of brain regions and hundreds of thousands of papers on how they're connected. Also, they were the first to sort through all the different nomenclatures and sub-literatures in the world to create a coherent database of the connections between different parts of the monkey brain.

So that's interesting, and eventually if you bring that kind of connectivity diagram together with the kind of simulation that they did, potentially you could get a large-scale simulation with more of the same structures and dynamics as a real animal's brain – but they haven't gotten there yet.

[Open Connectome](#) is another interesting project, at John Hopkins University, to mention in that regard. It's a little bit earlier stage than what Modha's team did with the monkey brain, but it's all Open Source. Their scientists upload connectivity data from different parts of the brain, and make open source tools where anyone can go online and help map out neurons, synapses and what's connecting to what in the data – and this could produce a much more fine-grained map of the connectivity structure. If something like that succeeds, then you could really make a large-scale brain simulation that does what the brain does – which is something that neither Markham nor Modha did in their simulations.

**SM Dambrot:** That kind of open-source project would have a significant benefit to a wide community of neuroscientists.

**Dr. Goertzel:** Yes – they want to go Web 2.0 with it: They want to not only have scientist upload their data, but also have people from around the world log on and help interpret the data. It’s interesting – there are some image processing tasks that people are good at but computers aren’t that good at. For example, with three-dimensional imaging data – the type of data that the John Hopkins researchers have uploaded – people can look at and see, “yes, there’s a neuron *there*, and it’s pointing to another neuron over here.” Current image processing tools, however, are quite weak with 3D data.

So right now, there’s a role for people to look at this 3D data and see what’s connected to what. Once AI is a little further advanced at 3D image processing tasks, the role of people will shift to correcting the AI’s mistakes, and then ultimately the AI could obsolete people – in part by leveraging the training data obtained from people’s image classification judgments made by using the Open Connectome web interface.

**SM Dambrot:** Would you consider this the next step in the progression of distributed processing – SETI@home, ProteinFolding@home, and so on?

**Dr. Goertzel:** In a sense – but those are using home computing power to do number crunching, whereas Open Connectome uses human brain power. It would be interesting if you could take a page from the [Google Image Labeler](#) that Luis von Ahn created at Carnegie-Mellon University – he made labeling images online into a game to make it fun for people to provide textual labels for images, but it’s a game with a purpose: the labeling then serves as AI training data. It’s not exactly *Name the Neuron* – the point is not to label a neuron but rather to identify it and where it’s connecting – but I think it could be approached in a similar way.

**SM Dambrot:** Another interesting topic from your talk yesterday was the use of virtual and gaming worlds to provide and AI with a space to

explore – specifically the block world.

**Dr. Goertzel:** In the AI project I'm currently doing with [Hong Kong Polytechnic University](#) (PolyU), the basic goal is to demonstrate OpenCog doing something in a videogame world which will be interesting to the game industry. At the end of this two-year project, which is jointly funded by the Hong Kong government and my company Novamente LLC, we want to create an OpenCog agent in a game through a partnership with a game company – to both generate money for ongoing research, and establish a way to set the AI up in communication with potentially millions of people around the world who would be the AI's teachers.

Then the question becomes: What type of game world should we use for our current prototype experiments? We've done some work before using a game platform called [Multiverse](#) in which the actor is a virtual dog that learns tricks – which is interesting as a platform for imitation reinforcement learning, but it's limited. We wanted something with more versatility but not so much that it would confuse our early-stage AI.

An AGI Preschool is a cool idea. I want to do it, but it's a bit much for right now – less in terms of the AI, which could probably handle it, but in terms of resources for game development. In a preschool you have a lot of things that are hard to simulate in a video game – a sandbox and Play-Doh, for example – so we settled on a game world modeled on the video game [Minecraft](#) because it's relatively simple from a game development perspective yet provides a lot of flexibility on terms of the AI interacting with the world. In Minecraft, everything in the world is made of small blocks, which can be used to build anything – a ladder, tower, or even a statue that looks like oneself. There's a lot of opportunity for flexibility and creativity, but because everything is made out of blocks you don't have to deal with scripting sand and other difficult objects, and you don't have to do as much artwork and

animation.

In short, we made this decision to both simplify the AI's job in terms of perception and action so it could focus more on cognition, learning, planning and construction, as well as to simplify game world construction – a world made of blocks is basically Democritus's model of the cosmos, on a larger scale.

Still, there are various decisions to make – in the physics of the game world, for example, you can build a very narrow tower of blocks but gravity doesn't make it fall down.

**SM Dambrot:** Adding realistic physics would give you the best of both: you'd have real-world constraints coupled with the simplicity of using repetitive units to construct objects.

**Dr. Goertzel:** That's right. And of course, in terms of transfer to a physical robot, you can give that robot blocks to play with in the robot lab. It transitions fairly well into building with wooden blocks, Lego blocks and so on. This natural transition path for the game world into robotics will probably be done in the Hong Kong project, which is focused on game AI.

**SM Dambrot:** You also discussed various types of memory in human cognition. Does AI memory conform to these?

**Dr. Goertzel:** Overall, my approach to AI is not based on neuroscience, primarily because I don't we know enough about neuroscience to drive AI design – and the neuroscientists I talk to tell me the same thing. It is inspired by cognitive psychology to a significant extent. The different types of memory I used to design OpenCog are pretty well established in Cognitive Psychology, in the sense that we seem to have different mechanisms with different response time characteristics for, say,



procedural knowledge versus semantic knowledge. If you dig into the neuroscience, there are many distinctions between these types of memory, in that various parts of the brain are differentially active during types of memory. For example, there's evidence that the cerebellum is involved during action sequences – the basal ganglia also come into it – even though they don't involve motor action. In spatial knowledge, there are complex interactions between the posterior parietal cortex, hippocampus, entorhinal cortex, and so forth. We're not at the stage where neuroscientists have a clear picture of how each of the different types of memory is implemented. So clearly there's the same biochemical and cellular mechanisms underlying different kinds of memory in the brain, and there's much overlap in terms of the brain regions and dynamics, as well as there being significant differences in which brain regions are involved, and in which neurotransmitters may be involved. The details are still unfolding.

If you look at what you can do on a computational neuroscience level now, you can do things like build a model of the hippocampus and medial temporal lobe, connect it to your model of the parietal cortex, and study how that implements spatial memory. The hippocampus and medial temporal lobe tend to deal more with allocentric coordinates (such as third-person top-down, or bird's-eye, views), while the parietal cortex tends to handle first-person egocentric views – but both are head- and eye-centric. Neuroscientists have different opinions about the brain's coordination of these different perspectives – and I've been doing some consulting in this direction through Novamente. However, to me this is a different pursuit than trying to build a human-level thinking machine, because the neuroscience is just too diverse, particular and unfinished.

**SM Dambrot:** Especially given the idea that AGI is ideally substrate-independent.

**Dr. Goertzel:** Substrate independence is an interesting notion, and as a



mathematician I would like to aspire to that – yet as an AGI designer I’m constantly pushed away from it. The OpenCog design now is not that substrate-independent – in fact, in many ways it’s customized to operation on a network of symmetric multiprocessor Von Neumann machines.

In the just-finished first draft of my new book *Building Better Minds*, the core mathematics is substrate-independent – for instance it would work on a massively-parallel MIMD machine, like the Connection Machine that Danny Hillis built at MIT starting in the 1980s – but on the other hand, there’s also a lot of content and code heavily tied to the particular hardware we’re currently using. For example, we have to write code to multithread among 16 processors (or however many processors our individual SMP machines have), and we then will have to write code to network many of multiprocessor machines together. That has a lot of consequences – for example, if you’re running on 1,000 machines, each with 100GB of RAM, you have issues of how to dynamically and adaptively partition knowledge among those machines. How do your logical inference control and procedure learning mechanisms make use of this clustered feature of your knowledge base?

Once you go in that direction you’re adapting your systems to a network of symmetric multiprocessor machines, which is an infrastructure that very different from a Connection Machine or human brain – so if you gave us a Connection Machine with a trillion processors, we could port our mathematical algorithms, but much of the code would have to be rewritten, as would the intermediate layer of algorithms that we use as a “glue” between the mathematics and the hardware.

In short, efficiency leads you away from substrate independence, so as an AGI designer you want to formulate your core cognitive algorithms and structures in a substrate-independent way. At least that’s my approach. On the other hand, you could take a different view: If you're

less of a mathematician and more of an engineer or biologist, then your approach could be to grow a mind out of the substrate, which is what happens with the human brain – evolution didn't start with abstract mathematics of thought that was then implemented on wetware.

**SM Dambrot:** This reminds me of our discussion a few minutes ago about the ways worlds and minds interact, in that the brain is tied in with the world in which it evolved.

**Dr. Goertzel:** The brain is part of the world – it's made of the same stuff as the world around it. It's more a matter of one part of the world co-evolving with another – and what we're doing with AGI right now is engineering, not evolution.

A long time ago – before I started seriously working on AGI – I had the same thought many others have: Why not evolve a brain by implementing an artificial ecosystem across the Internet, set some artificial chemistry and biology in motion, and let the AGI emerge from the digital primordial soup. The obvious conclusion you come to after a while, yes, that's really cool – but the ecosystem has many more molecules than any one brain, and that's going to require orders of magnitude more computing power than does any individual brain, so it's probably not the best approach to take.

**SM Dambrot:** Since we're at the Humanity+ Transhumanism Conference, my last question is about the connection between your work in AGI and Transhumanism.

**Dr. Goertzel:** From a certain standpoint, working on an AGI is a purely technical and engineering pursuit which could be done by a lot of people – such me and five or ten other guys locked in a basement somewhere, just coding our hearts out all day. On the other hand, that's not really the way things are going – we're developing our AGI in an Open Source

project with people around the world, trying to recruit new programmers, and with funding that so far has largely been based for vertical market applications, not just for pure research. Therefore, in practice – since our development of AGI is distributed around the world and couple with business, universities, and various other entities within the world – there’s been a fair amount of interoperation between the AGI outreach and the Transhumanism outreach that I’ve been doing.

As an example, our AGI project in Hong Kong Polytechnic University – where we’re developing OpenCog for video games – involves Gino Yu, who runs the lab, but who with me is also organizing the Humanity+ Hong King conference on December 3-4, 2011. Through that conference, we’ll get Hong Kong technology and business people attending, potentially leading to connection for more OpenCog commercial projects or university collaboration, in turn potentially leading to funding that will feed OpenCog development.

There’s a lot of cross-pollination scientifically as well: The OpenCog work is integrating many different AI tools, one of which is *machine learning* – a particular AI discipline based on learning by example that could itself be integrated with probabilistic reasoning, analogic inference and generalization. I’m using machine learning in my bioinformatics work to analyze genetics data – and in that bioinformatics work I’m collaborating with Genescient, accompany whose founding Chief Scientist was Michael Rose who I met at the Transhumanism-related Immortality Conference in 2005.

What I’d like to do in the next couple of years, among many other things, is to use OpenCog for the genetics work by pulling in probabilistic reasoning and concept learning so that we’re not just doing machine learning, but are also doing some AGI-type cognition about that bioinformatics data. That would be a case of OpenCog integrating more advanced technology into a bioinformatics project or engineered life

extension, which was founded through a connection made at another Futurist conference. At the moment, it's all one big social and intellectual network, rather than being siloed into AGI, Transhumanism, and so on. To a large extent, that's my own personal approach – there are certainly very solid AGI researchers who have no connection with the Transhumanist community, and of course there are Transhumanists thinking about AGI who have no connection with AGI research. I'm always interested in connecting things together – my main focus in life is making intellectual progress on scientific issues, but I spend a certain percentage of my time pulling people, social networks and ideas together, which I think is also valuable.

As a final example, at the AGI '11 Conference – a technical AGI conference which will be held at the Google campus in Mountain View, California – we'll have a Future of AGI Workshop before the conference, which should attract Transhumanists who wouldn't necessarily attend the technical meeting. Pulling the community together like this can have a lot of impact – some Transhumanists may be involved in practical projects that could benefit from AGI technology, others or their friends and associates may have a technical background and so might want to get involved with AGI work, and of course meeting and talking with real AGI theorists may help them speculate about the future about ways that are better grounded than might otherwise have been.

**SM Dambrot:** If you would, please take a final moment to give us additional details about the AGI and Transhumanist conferences later this year, as well as when we might expect your upcoming books.

**Dr. Goertzel:** [AGI 2011](#), to be held in Mountain View on August 3-6, is in large part a technical and scientific conference for those involved in Artificial General Intelligence, but the pre-conference workshop, as well as the Keynotes and demo sessions, will be interesting to everyone – so I

encourage you to register soon, as there's a cap on attendance of some 200 attendees due to the size of the venue at Google.

The [Humanity+ @ Hong Kong Conference](#) will be held on December 3-4, 2011, at Hong Kong Polytechnic University's Chiang Chen Studio Theatre. It should be very interesting in terms of bringing in scientists and futurists from mainland China who don't circulate much in the world-at-large or intersect with their Western counterparts – so I'm psyched about the cross-cultural admixture there.

In terms of my technical AGI book, *Building Better Minds*, its release date of course depends on the publisher, but my guess would be at the late 2011 or early 2012. I'm also working on an AGI trade book, tentatively titled *Faster Than You Think*, which should also come out in 2012.

**SM Dambrot:** Thank you so much, Dr. Goertzel.

**Dr. Goertzel:** Thank you for the interview.

*This is the second part of a two-part article. The first part is available at <http://www.physorg.com/news/2011-06-dr-ben-goertzel-artificial-intelligence.html>*

Copyright 2011 PhysOrg.com.

All rights reserved. This material may not be published, broadcast, rewritten or redistributed in whole or part without the express written permission of PhysOrg.com.

Citation: Interview: Dr. Ben Goertzel on Artificial General Intelligence, Transhumanism and Open Source (Part 2/2) (2011, June 13) retrieved 27 April 2024 from [http://phys.org/news/2011-06-dr-ben-goertzel-artificial-intelligence\\_1.html](http://phys.org/news/2011-06-dr-ben-goertzel-artificial-intelligence_1.html)

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.