Creating a Robot Culture: An Interview with Luc Steels

Tyrus L. Manuel

May/June 2003

Histories and Futures regularly features two types of people: visionaries whose past work helped to shape the field and those who seek to take AI to new levels into the future. Luc Steels fits into both of these categories.

From his work with natural language and robot behavior— which became especially visible with Sony's Aibo—to his new focus on building robot communities and studying robot interactions, he has continually sought to redefine the AI field. In this interview, he shares his opinions with *IEEE Intelligent Systems* on AI and how the field is evolving.

How did you get your start? Was there some moment in your life that sparked your interest in AI?

My first field was linguistics. In the early seventies, this field was very exciting with the rise of generative grammar and Montague grammar. But then I discovered, almost by accident, the power of computers and was enormously impressed by experiments such as Shrdlu by Terry Winograd around 1972. I learned that it was possible to test theories of language processing through computer simulations. The intuitions that I and most other people (including linguists and psycholinguists) had concerning language processing at that time were very naive and simply could not be made to work. So I decided to study computer science at MIT and have become totally immersed in artificial intelligence ever since.

Does language hold the key to helping AI better understand and relate to its environment?

Intelligence comes in many forms. Some people consider an ant colony or genetic evolution to be intelligent. In those cases, there is no language—cognition isn't even involved. But for human-like intelligence, language is indeed essential. It is one of the main ways in which human beings bootstrap themselves into cognition and culture. So it is not just a simple interface to complex thought processes, it actually shapes and enables thought. Through language, we coordinate ways to view the world.

A great deal of research seems to be moving forward based simply on a need to create a sentient machine or to be able to beat the Turing Test. Do you agree?

In the history of AI, almost no one has ever seriously worked on the Turing test, contrary to what many journalists, science fiction authors, and philosophers seem to think. There has hardly been a paper published in a serious AI conference or AI journal in which the author claims that he or she is working on the Turing test or has solved it. There are no serious projects to my knowledge nor has any serious funding agency ever provided money for this. The Turing test is not the challenge that AI as a field is trying to solve. It would be like requiring aircraft designers to try and build replicas of birds that cannot be distinguished from real birds, instead of seriously studying aerodynamics or building airplanes that can carry cargo (and do not flap their wings nor have feathers).

Instead, there are two basic motivations for work in AI. The first motivation is purely pragmatic. Many researchers try to find useful algorithms—such as for data mining, visual image processing, and situation awareness—that have a wide range of applications. They are challenged by the intrinsic difficulties of certain problems or the potential benefits of new applications.

The second motivation is scientific. Computers and robots are used as experimental platforms for investigating issues about intelligence. Researchers who are motivated in this way, and I am one of them, try to make contributions to biology or the cognitive sciences. Although there was a lot of resistance against the synthetic methodology initially, it is now widely accepted, and AI has had an enormous impact on how we think today about the brain and the mechanisms underlying cognitive behavior.

So you believe that eventually the Turing test will fade and we will no longer seek to have robots exactly duplicate humanity. What will eventually cause this change?

Most people in the field have known for a long time that mimicking exclusively human intelligence is not a good path for progress, neither from the viewpoint of building practical applications nor from the viewpoint of progressing in the scientific understanding of intelligence. Just like some biologists learn more about life from studying a simple *E.coli* bacteria, there is great value in studying aspects of intelligence in much simpler forms.

Will robots ever become the planet's dominant life form?

I don't believe for a moment that robots will dominate the earth. These ideas are in the fantasy world of science fiction authors or people who have no clue of the enormous complexity of human intelligence. There is nobody in AI that seriously works on projects in this line. Why would we ever build machines that could destroy us? Robots can either be very useful—but then they are specialized and usually not very autonomous or cognitive—or they can be entertaining, like Sony's pet robots. When entertaining, robots become more of a medium like opera or puppet theatre, creating artificial worlds that we enter to enjoy ourselves.

Why have you recently gone beyond your focus on language to stress the importance of interaction in robotic learning?

Before I started to work (again) on language in 1995, I focused on behavior-based robotics. In fact, my recent work on language evolution is a continuation of work in behavior-based robotics and artificial life, in an attempt to try to bootstrap cognitive intelligence in robots. The new thing, compared to earlier work in language processing, is that I want to investigate how language can be grounded in the real world and how language can emerge in a community of artificial agents.

Your new theories diverge from the common concept that AI breakthroughs must be achieved by building more advanced machines. Your approach runs parallel with how humans learn and develop, so why has much of the AI community met your ideas with resistance and skepticism?

What we need today (and I think most people in AI would agree) is not really more powerful or novel hardware but new ideas. New ideas will always be received with skepticism. If there is no resistance, the idea is simply not revolutionary enough.

In the earliest phases of AI, there was a greater openness and much more variety and freedom of thinking than today. Many

AI researchers are too focused on shortterm applications. However, we must push hard to explore new radical avenues if the field wants to make any progress. The work that we have been doing lately on language evolution is getting an enormously positive response from biologists, linguists, anthropologists, and other disciplines interested in questioning the origins of languages and meaning. There is a rapidly growing activity attempting to get communities of robots or agents to evolve their own languages and ontologies.

On the basis of your new theories, do you envision "classrooms" where separate robots learn certain things and then are all brought together to share their knowledge with each other?

Classrooms might not be the best environment for learning (not even for human children). Robots need to be in the world, interacting freely with an open-ended environment, autonomously seeking new challenges and opportunities for cognitive growth and interacting with other robots and human beings. It is only by an extended cognitive development that we can ever hope to get closer to human-level intelligence. Moreover, it is definitely true that robots can be very effective to bootstrap each other, partly because if they start at the same level, their competences are naturally scaffolded and they push each other up.

Does this also mean that robots can do a better job of teaching fellow robots than humans can?

AI designers have been criticized (sometimes rightly) for putting in too much by hand. For example, an expert system in which rules are programmed that were extracted from a human expert is not considered to be autonomously intelligent; it is just a way to capture and automate something that humans have already figured out. The same thing is true with language. If we put in the rules of English, we do not explain how an intelligent being could learn or invent language rules. Language is a moving target. New sounds, new meanings, extensions of word meaning, and new grammatical constructions appear all the time, and a language user creatively innovates. This is the process we need to understand.

If we let a robot community develop its own culture, ways of viewing the world, and ways of communicating about it, the artificial communication and representation systems the robots develop might have language-like features, but they obviously will never be equal to existing human languages. There are too many contingencies that shaped a particular language such as English, and the robots might have totally different ways to sense the world or might need to communicate about topics completely alien to us. What interests me, however, is that they developed these communication systems themselves.

Your current research stresses a humanlike learning model, yet you don't like to measure robot evolution by human standards, correct?

Human intelligence is obviously the most magnificent manifestation of intelligence in the natural world. So it is quite natural that this remains a source of inspiration. But we cannot be only human-centered in our evaluation criteria. For example, if a community of robots develops a shared communication system with human language-like features, we should not measure this against English. Instead, we should use more abstract criteria, asking, for example, if there is an emergent grammatical structure in the language, how fast can new features propagate in the group, and so forth.

Do you have any recommendations on how robotic evolution should be measured?

The evolution of robots will be a matter of slow progress, simply because so many technologies must come together. For example, battery technology or motors for joints is as critical as getting enough computer memory or good algorithms for vision and speech. One thing we could try to establish in the field is a road map, comparable to Piaget's account of development in children (although we would have to add steps in which the group becomes more effective as well). For some subareas of AI, such roadmaps exist, but making such a roadmap could be a very good exercise.

Finally, what do you think the future holds for AI?

Current work in AI already focuses on many forms of intelligence not at all in the line of the Turing test—for example, work on embedding intelligence in ubiquitous computing environments. AI should work more like a scientific field in which systematic experiments are done and redone to examine particular hypotheses, and in which there is theoretical research to find out why certain things work and others do not. It is only through application of the scientific method that there can be a progressive accumulation of knowledge. AI has grown out of a pragmatic tinkering approach, which

worked well in the early stages but is not a good method for building a systematic theory of intelligence. AI should also be connected more to other scientific fields, particularly biology, so that ideas can flow more in all directions and experimental techniques can be carried over.

The problems AI is addressing are vast and extremely difficult. AI is more difficult than physics or biology because it is dealing with complex adaptive systems with many layers of complexity. The future is bright, both in terms of scientific challenges and potential application areas. But we should not pretend that all problems are solved or human-level intelligence is near. Intelligence is like life or the cosmos; it is such a deep phenomenon that we will still be trying to understand it many centuries from now.

Luc Steels is a professor of computer science at the University of Brussels (VUB), director of the VUB Artificial Intelligence Laboratory, and director of the Sony Computer Science Laboratory in Paris. His scientific research interests cover the entire AI field, including natural language, vision, robot behavior, learning, cognitive architecture, and knowledge representation. His current research focuses on dialogues for

humanoid robots and fundamental research into the origins of language and embodied meaning. Contact him at steels@arti.vub.ac.be; http://arti.vub.ac.be/ ~steels.

Tyrus L. Manuel
Creating a Robot Culture: An Interview with Luc Steels
May/June 2003

<https://ai.vub.ac.be/sites/default/files/manuel-03a.pdf>
Features Editor: Rebecca L. Deuel: rdeuel@computer.org
IEEE Computer Society