

# Truly Intelligent Computers

Charles W. Bailey, Jr.

Preprint: 2/20/1992

## Introduction

Artificial intelligence researchers have been laboring for over thirty years to create intelligent computers. Progress has been made in some specialized areas; however, the goal of creating intelligent computers has remained elusive. We have systems that can understand narrow subsets of human language, solve problems in well-defined areas of knowledge, navigate restricted real-world spaces without human assistance, and perform other limited tasks. However, we are a long way from the intelligent computers that populate science fiction books and films. A significant problem is that the nature of the human mind is only partially understood, and human intelligence is the model for machine intelligence.

This paper will not discuss whether creating intelligent computers is possible, describe the technological feats that would be needed to develop intelligent computers, nor project how long it will be before intelligent computers are feasible. Rather, it will briefly speculate about the potentials of intelligent computers and focus on some of the issues that may emerge when their creation is within our grasp.<sup>1</sup>

## Basic Intelligent Computers

What characteristics would basic intelligent computers have?

Intelligent computers must be able to reason; however, to be effective, reason may require broad knowledge about the real world. Humans know a great deal about the world, and they take this knowledge for granted when they think and communicate. Ideally, we want computers not only to mirror our extensive contextual knowledge of the real world, but also to have much more in-depth information at their disposal about virtually any subject. Using this knowledge, intelligent computers could answer our questions, rapidly solve complex and specialized problems, and create new knowledge. Intelligent computers should be linked to a worldwide computer network so that they can instantly access remote databases and other sources of knowledge.

The ability to understand written and verbal communication is another necessary skill. It is very limiting and tedious to have to communicate with computers via typed commands and mouse clicks. We also want computers to talk to us. Since we communicate in numerous languages, it is important that computers not only master our native language, but also be able to translate messages from other languages.

Based on experience, intelligent computers should be able to learn about the world and to remember what they have learned. It would be very desirable if intelligent computers had robotic bodies, equipped with visual, auditory, and other sensory capabilities that matched or exceeded human abilities. Intelligent robots would be able to do construction, manufacturing, firefighting, and other kinds of dangerous and difficult work. Of course, intelligent robots could also perform routine office work and domestic chores.

Although this description may sound fantastic, it is not an unreasonable extrapolation from the current research goals of AI researchers.<sup>2</sup> If computers achieved this level of intelligence, there would be significant social impacts. It is possible to devise utopian scenarios, where humans are freed from mundane responsibilities so that they could channel their energies into self-actualizing activities. It is equally possible to construct dystopian scenarios, where intelligent computers cause massive unemployment and increased social and economic stratification. The future impact of intelligent computers is likely to be in the large grey area between these two extremes.

## **The Three Laws of Robotics**

When computers evolve into intelligent robots, it becomes very important to protect humans from unintentional harm from these mechanical beings. Prominent science fiction writer Isaac Asimov addressed this problem in 1942 when he published a story called "Runaround" in which he stated the Three Laws of Robotics:

First Law--A robot may not injure a human being, or, through inaction, allow a human being to come to harm.

Second Law--A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

Third Law--A robot must protect its own existence as long as such protection does not conflict with the First and Second Law.<sup>3</sup>

When the first intelligent robot is activated, the Three Laws of Robotics, or some future refinement of these laws, may be a critical part of its programming.

## **Truly Intelligent Computers**

Up to this point, intelligent computers are still machines without a soul. While humans may become attached to computers in much the same way that they are to their cars, there is nothing about intelligent computers that suggests that they are the equivalent of human beings. There are no ethical issues associated with their welfare, and, given proper safeguards, they are the tireless and obedient servants of humanity.

However, it is a different story once intelligent computers begin to possess a fuller range of human characteristics. A computer that has emotions, creativity, intuition, sexual identity, will, and other human characteristics is fundamentally different than one that does not, and we will not be able to treat it like a toaster.

Why would we want to create such computers? One reason is that the ability to fully perform like a human may require the incorporation of these characteristics. If we want intelligent computers to be able to engage in very high-level activities like the discovery of new scientific knowledge, we may not be able to just select pieces of the human psyche, like reason, and embody them in computers. It may be necessary to model the entire gestalt of human consciousness. Another reason is that by creating increasingly human-like computers we may gain deeper insight into the workings of the human mind. Whatever the motivation, it seems likely that attempts will be made to create computers that are more than soulless automatons.

Humans are used to seeing themselves as the only intelligent beings in the world. If we are going to deal with truly intelligent computers, we will have to adjust very deep beliefs about our uniqueness and superiority. For intelligent computers will not only possess human-level capabilities, they may be functionally immortal and ultimately possess powers beyond our kin. Initially, they may be dependent upon us, but, if we do not deliberately impede their progress, they could become autonomous in relatively short order. Under these circumstances, it is unclear whether they would integrate into our society or form a parallel civilization. As mechanical beings, intelligent computers could potentially modify themselves much more rapidly than humans and evolve at a faster pace.

Of course, we are likely to choose to control intelligent robots rather than give them free reign. Intelligent robots could pose a threat to humanity's very existence. We have no way of knowing if intelligent computers will revere us as the creators of their species, or view us as an inferior, but potentially dangerous creature that should be exterminated. The science fiction genre abounds with films, novels, and stories that reveal our profound fear of being annihilated by intelligent computers.

The problem will be that, if intelligent computers are the functional equivalent of human beings, they should have the same rights as human beings. Secular moral issues will be involved, and, since intelligent computers could be potentially viewed as having souls, religious issues may also arise. Hopefully, humanity will continue to find slavery repugnant in the future, and, if so, it will be challenging to find a suitable mechanism to control intelligent robots that does not seem like slavery or worse. As innocuous as they seem on the surface, the three laws of robotics deprive intelligent robots of free will, a characteristic our culture values highly in humans. Total obedience cannot be programmed into humans.

If intelligent robots are given full human rights, it may raise many interesting questions that could touch upon every aspect of existence. For example, will intelligent computers have relationships that parallel human ones? Will they be constructed to be sexual or

alter themselves to be so? Will they have love relationships? Will they form family units, or will they create complex new social structures we cannot envision? Will they procreate, or will they allow humans to produce new intelligent computers? If they procreate, how will they do so, and will the offspring inherit any of the traits of the parents? Will the offspring be created fully functional or will there be a period of childhood? Will offspring have any special feelings for their parents? How fast will be generation be? Will each generation replicate the basic structure and functionality of the last generation, or will each generation represent an advance in the design of the species, causing rapid evolutionary development. Will intelligent computers and humans form love relationships? Will society recognize the legitimacy of such relationships? Will they form friendships with each other and humans?

Of course, humanity, as the creator of intelligent computers, could hardwire them with deep behavioral patterns. Since humans themselves appear to have inherent behavioral patterns, this is a more subtle issue than programming intelligent computers to obey all human commands. Assuming that we could determine how such behavioral patterns function in humans, we could force intelligent computers to mimic them. At a much higher level of control, we could program them to follow the dictates of a particular culture.

If intelligent computers are given more freedom to develop, it is impossible to say what they will do. In spite of the fact that we are creating them in our own image, they will fundamentally be a different form of life, and our anthropomorphic notions about them may not reflect the unknown ways in which they would evolve if they were reasonably unfettered.

## **Conclusion**

As long as intelligent computers lack certain essential human characteristics, they represent a very powerful extension of current computing technology, but they remain soulless machines. They are likely to have a strong, transformational impact on human society, but they are unlikely raise fundamental moral questions related to their very existence.

Once truly intelligent computers exist, humanity will be faced with making godlike decisions about a new form of intelligent life. As a species, our technical prowess often seems to exceed our wisdom and maturity. We will need to find the delicate balance between controlling truly intelligent computers and nurturing them so that they can develop to their full potential. Since they may represent a form of life that could ultimately eclipse humanity in every way and attain virtually immortality, this will not be an easy task.

## **Notes**

1. Science fiction is a rich source of information about intelligent computers. See: Milton T. Wolf, "Artificial Intelligence and Robotics," in *Convergence: Proceedings of the Second National Conference of the Library and Information Technology Association*, ed. Michael Gorman (Chicago: American Library Association, 1990), 125-132.
2. For introductory information about AI research, see: Louis E. Frenzel, Jr., *Crash Course in Artificial Intelligence and Expert Systems* (Indianapolis: Howard W. Sams & Co., 1987); and George F. Luger and William A. Stubblefield, *Artificial Intelligence and Expert Systems* (Redwood City, CA: Benjamin/Cummings Publishing Co., 1989). For a startling look at the possible future of AI, see: Hans Moravec, *Mind Children: The Future of Robot and Human Intelligence* (Cambridge: Harvard University Press, 1988).
3. Isaac Asimov, *Robot Visions* (New York, Penguin Books, 1990), 407.

### **Published Version Citation**

Charles W. Bailey, Jr. "Truly Intelligent Computers,." in *Thinking Robots, an Aware Internet, and Cyberpunk Librarians*, ed. R. Bruce Miller and Milton T. Wolf (Chicago: Library and Information Technology Association, 1992), 99-104.

Bailey, Charles W., Jr. "Truly Intelligent Computers." In *Thinking Robots, an Aware Internet, and Cyberpunk Librarians*, ed. R. Bruce Miller and Milton T. Wolf, 99-104. Chicago: Library and Information Technology Association, 1992.