3-22-2010

Artificial Intelligence: Soon to be the world's greatest intelligence, or just a wild dream?

Edward R. Kollett
Johnson & Wales University - Providence, edk540@students.jwu.edu

Artificial Intelligence:

Soon to be the world's greatest intelligence, or just a wild dream?

Edward Kollett
Johnson & Wales University
Honors Program
2009

Artificial Intelligence is a term, coined by John McCarthy in 1956, that “in its broadest sense would indicate the ability of an artifact to perform the same kinds of functions that characterize human thought processes” (“Artificial Intelligence”). Typically, it is used today to refer to a computer program that is trying to simulate the human brain, or at least parts of it. Attempts to recreate the human brain have been a goal of mankind for centuries, but only recently, as computers have become as powerful as they are now, does the goal of a fully automated robot with human intelligence and emotional capabilities seems to be within reach. However, this paper will attempt to show that this is not actually true, and that since the founding of the Artificial Intelligence community, more steps backwards have been taken than steps forward. In addition, this paper will also attempt to show that even if such a robot could be created, it would be less efficient and infinitely more dangerous than the alternatives, many of which already exist today. In order to prove these two statements, the history of Artificial Intelligence as a field will be explored, including its successes and its failings, as well as its current day dilemmas. In addition, modern day applications of machine intelligence will be displayed, showing how much more efficient they are than an “all-purpose” super-intelligence. Finally, to prove that an “all-purpose” super-intelligence is more dangerous than it is useful, examples of moral dilemmas, social difficulties, and sources of other potential problems will be explored.

Perhaps the most important question about artificial intelligence to an average person would be why we’re putting so much time, money, and brain power into it. What can it do for us? It is theorized that an AI that meets or exceeds human intelligence could potentially become a better researcher than a human, and if that does in fact happen, than

AI could help us by greatly accelerating any and all research in science, medicine, and yes, even artificial intelligence (Kurzweil). In addition, some AI researchers believe that a super-intelligent machine could very well be *the last invention humans ever need to make* (Bostrom), as, from that point on, the machine could make any new inventions faster, cheaper, and with less risk. Many believe that an artificial super-intelligence could easily solve problems that baffle humans such as nanotechnology, advanced space travel, nanomedicine with the capability to allow humans to live indefinitely, and, by altering its own source code, creating more efficient and more powerful artificial intelligence (Kurzweil). Of course, this last point should sound a little unnerving, as the idea of an intelligent machine modifying its own code could lead to a machine that stops caring about human life, and then as is easily imaginable, goes on a rampage. This is perhaps the biggest risk associated with artificial intelligence, and one that is certainly hard to address if developers are going to allow the machine to self-modify, one of the most important aspects of intelligence. But before going any deeper, the history of artificial intelligence deserves some attention.

The history of “Modern” Artificial intelligence dates back to about 1950, the year when the famous mathematician Alan Turing published his work “Computing Machinery and Intelligence”. In this book, Turing lays out the groundwork to his ‘Turing Test’, which would come to be a famous, if quaint, way to measure the intelligence of a machine. In the Turing Test, a person would ask questions, which would be answered by a human and the machine in question. In the experiment, the machine and the person were kept separated from the person asking the questions. If the person was unable to determine which respondent was the human and which was the machine simply from the answers, then the machine could be considered intelligent ("Artificial Intelligence"). Despite its almost quaint design, the Turing Test is still considered to be a great benchmark test to any prospective artificial intelligence system. Additionally in 1950, American mathematician Claude Shannon, known for his expertise in Information Theory, published two papers about the possibility of one day creating a machine capable of playing chess ("Artificial Intelligence"). Eventually that did happen, and the machine Deep Blue became famous for its chess-playing ability.

A few years later, in 1956, a workshop at Dartmouth College was created, based on "the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" ("Artificial Intelligence"). Basically, it was a meeting of the greatest technological minds of the time to discuss the idea and process of systematizing the process of human thought and learning so that a machine would be able to replicate it. Some of the noteworthy participants included mathematician Claude Shannon, John McCarthy (who, as stated above, coined the term artificial intelligence), Allen Newell, Herbert A. Simon (who won a Nobel Prize in 1978 for his work in Economic Sciences), and Marvin Minsky, who at one time described the human brain as a "computer made of meat" ("Artificial Intelligence"). Out of this meeting the field of artificial intelligence was officially born, and these early researchers were very optimistic about recreating human intelligence on computers. Many assumed that most if not all the problems with it would be substantially solved within a generation or two. This is also the first time when researchers in the field of artificial intelligence overestimate their abilities and underestimate the complexity of the human mind, behavior that will almost become synonymous with the field in later years (Kassan).

Although the Dartmouth College participants obviously underestimated the challenge of creating artificial intelligence, it would be unreasonable not to mention that there have been significant breakthroughs and achievements concerning AI. For example, in 1997, a computer program called Deep Blue was able to defeat the world chess champion Garry Kasparov in a Grandmaster-style chess match, which is a best of six games style match. In 1996, Kasparov had been able to beat a less powerful form of Deep Blue 4-2, but in their rematch a year later the upgraded Deep Blue program came out on top, beating Kasparov 3.5-2.5 (tie games were counted as one-half of a point) ("Artificial Intelligence Beats...”). This victory for Deep Blue was considered a substantial breakthrough for artificial intelligence due to the fact that chess has long been considered a game that took great genius to be able to play well. The game focuses on making creative decisions to emerge victorious, and to that point in time, no one had believed that computers would ever be more creative than humans. However, many will still argue that Deep Blue did not show creativity in its victory, as it merely calculated the best move after comparing hundreds of millions of possibilities. Deep Blue as a machine contained over 500 microprocessors, and was capable of analyzing 200 million moves per second. However, despite the fact that Deep Blue merely chose the best moves it could find, AI researchers argue that its creativity was in choosing which moves to check. As an example, in one match, Deep Blue chose not to take a particular move that many chess experts, including Kasparov, had expected it to take. This surprise move, which ended up being decisive in the machine’s eventual victory, was shown in its printout to be the result of checking how useful the move would have been 11 turns later. The fact that it had done such calculating on a move no human expected it to take is what many believe show that machines can, in fact, be as creative or even more so than humans. To sum it up, an editorial written by Drew McDermott, a computer science professor at Yale University, believes that, essentially, a master of chess is able to play so well because they have memorized thousands of positions and the best moves to use when in those positions. However, they don't consciously make these decisions, the answer just 'comes to them'. But in reality, if the human mind is subconsciously recalling the best move from the thousands it knows (which we would consider 'intelligent'), whereas a computer is checking the position against some database, if they each arrive at the same answer, they couldn't the two methods be considered 'intelligent' – especially since some could argue the human mind is just checking its own 'database' subconsciously? (“Artificial Intelligence Beats...”).

Another significant achievement in the field of artificial intelligence is Chinook, a computer program that was designed to play checkers. Work on the program began in 1989, and although it lost to Marion Tinsley, heralded as the world’s greatest checkers player, in 1992, its creators are now confident that no one will ever be able to beat it again. The reason is that the researchers believe that they have ‘solved’ the game, so that even a mistake-free player would only be able to tie the machine. This accomplishment is considered great because even checkers, which appears to be a simple game, takes great intelligence to play well. It is estimated that there are about 500-billion-billion ($5x10^{20}$) possible positions. The ability of the computer (or computers, as Chinook has been run on up to 50 computers at the same time) to calculate the best possible move to either win or prevent itself from losing is considered to be an indicator of how intelligent a computer could become in the future (Schaeffer et al).

Computers today are far more powerful than they were back when these great AI programs were written. Yet the true goal of artificial intelligence, a machine that can fully mimic a human being's intelligence, has not been created. One would certainly be justified in wondering why. After all, if such powerful programs could be written on a 'primitive' machine, surely we should now have the power to create a program equal to the human mind. In order to illustrate the differences in computing power from today and back then, let's compare the amount of Random Access Memory (RAM) and the processor speed (two aspects of a computer which are good indicators of its power) of a $1000 computer in 1997 and one you can purchase today. in 1997, a computer that cost $1000 had 16 megabytes (MB) of RAM and a processor that ran at 120-megahertz (Manes). Today, the same $1000 will purchase a computer, (the following was built at dell.com) with 4 gigabytes (GB, one GB is 1000 MB) of RAM and a dual core processor (each processor can act independently, essentially, you have two) that runs at 3.33-gigahertz (≈ 3330 -megahertz). In essence, you have a computer that has 250 times as much RAM (which means it has 250 times the memory capacity) and has two processors that each run almost 28 times as fast (its as though it had two brains that each think 28 times faster than the older one). So again, considering how powerful computers are now as compared to that "golden age" of AI, why are we still unable to replicate the human mind on a computer? The largest single reason why we do not have intelligent robots is because no one knows *what* to build, much less how to build it (Kassan). Some prefer to argue the point of Moore's Law in this case – which states that every 18 months, the processing power of computers will double (Malcolm). This has been shown to be more or less true. The argument is that since computers are rapidly becoming more *powerful*, they must also must be getting more *intelligent* – and this is not true. Computers, no matter how powerful, still “lack the general underpinning of "common sense" which we have. They are also incapable of the widely general, insightful, and rapid learning which characterises human students” (Malcolm). No amount of computing power can change that – in order to fix this, researchers will need to be able to figure out what ‘common sense’ and ‘general, insightful thought’ are, and how to program them – something unlikely to keep up with Moore’s Law.

In addition, certain other aspects of human intelligence, such as the ability to perform math or the ability construct a sentence, lend themselves well to programming (“Artificial Intelligence”). It is common knowledge that computers are far better than any human will ever be at performing mathematical computations, and most people are familiar with programs like Microsoft Word that tell you when you have misspelled a word or (occasionally) when you aren’t using proper grammar. However, it is very difficult to program other parts of human activity (“Artificial Intelligence”). For example, it would be difficult to program the computer to interpret non-verbal messages from a person when talking to it, and a computer likely cannot comprehend a concept like ‘what a person meant’ as opposed to ‘what they said’ – as such, Microsoft Word is unlikely to catch a mistake where you use the wrong (but grammatically correct) word. In essence, a computer is good at things like math and language syntax. Meaning and abstract concepts, however, are difficult for a computer to grasp, mainly because it is difficult for the programmers to explain it in code.

There is an attempt to overcome this challenge, however. Known as Fuzzy Logic, this new type of information processing attempts to make computers think in a more human way by adding shades of gray in between the computer's 1's and 0's. Dating back to the pioneering work by Lotfi A. Zadeh in 1965, Fuzzy Logic has become an important part of computing, and has become "a profitable tool for the controlling of subway systems and complex industrial processes, as well as for household and entertainment electronics, diagnosis systems and other expert systems" (Bauer, Nouak & Winkler). Fuzzy Logic works by replacing the simplistic true/false evaluations of a conventional computer with the full range of possible values from 0 to 1, with 0 representing an *absolute falsehood* and 1 an *absolute truth*, with everything in between being *somewhat true, somewhat false, very true*, etc. Essentially, it is an attempt "to apply a more human-like way of thinking in the programming of computers" (Bauer et al).

Fuzzy Logic represents perhaps the most useful application of artificial intelligence (besides perhaps the computer gaming industry), and certainly represents the most useful application in the realm of business. To understand more precisely how it works, consider this example given by Bauer and his colleagues. To categorize people as either young or not young, one could separate the groups based on a particular age, for example 20. But it makes little sense when you think about it: "why is somebody on his 20th birthday *young* and right on the next day *not young?"* (Bauer et al). This is a question we cannot resolve simply by changing the age. The problem arises because in this grouping someone either *is* or *is* not young, with no intermediate possibilities. Fuzzy Logic attempts to resolve this. If everyone up to 20 was instead considered 100% young, and, for example, everyone after 30 was 0% young, with everyone in between being somewhat young (in this example, 25 would be 50% young), there is a much more human result. People can be somewhat young, king of young, or not very young at all. This exact same method could be used to group people as being old or not. At 20 or younger, they are 0% old. By 30, they are considered 100% old, and all the expected intermediate values are present as well. Therefore, a person at 25 would be 50% young and 50% old. At 27, a person would be 30% young and 70% old. This allows computers to think in a more human way, with a range of probabilities for each possibility. Bauer and his colleagues list a few basic situations in which Fuzzy Logic can be used with great success:

(1)for very complex processes, when there is no simple mathematical model
(2)for highly nonlinear processes
(3)if the processing of (linguistically formulated) expert knowledge is to be performed

Though they also recommend staying away from Fuzzy Logic in some situations. Particularly, these:

(1)conventional control theory yields a satisfying result
(2)an easily solvable and adequate mathematical model already exists
(3)the problem is not solvable

As already stated above, Fuzzy Logic has been used in several different realms of business, typically to simplify a form of control. It has been applied in the “Automatic control of dam gates for hydroelectric-power plants”, for “Simplified control of robots”, to prevent “unwanted temperature fluctuations in air-conditioning systems”, and to improve “fuel-consumption for automobiles”, among many others (Bauer et al). It is easy to foresee Fuzzy Logic becoming even more prevalent in our lives as our understanding of it and our ability to apply it improve. It will likely become a dominating force behind any future artificial intelligence systems. After all, it allows the computer to think more like us.

Despite the advancements in fuzzy logic, problems that concern vague or imprecise qualities, such as those humans face every day, still pose a huge problem for AI researchers. Some believe that a more effective way to create artificial intelligence would be to create a machine that mimics the human mind as closely as possible. First proposed in 1943 by Warren McCullough and Walter Pitts at the University of Illinois, their idea only just beginning to seem possible (Fox). For example, a robot called Darwin VII was created by Jeff Krichmar and colleagues created at the Neurosciences Institute (NSI) in La Jolla, California. The robot had 20,000 brain cells roughly based on rat and ape brains (though far less complex), and was able to “learn” to find electricity-producing blocks (which it “liked”) after only half-an-hour of being turned on (Fox). Fox believes that these robots, with minds of silicon and bodies of steel, will be more successful than today's artificial intelligence systems. He believes that they will be able to use their artificial senses to 'learn', like humans, in this way becoming adept at handling the ever-changing situations and stimuli of our world. He envisions them one day being able to perform the “mundane household chores we'd rather not do”, driving kids to school, exploring Mars (or other uninhabitable places) or even run a nuclear power plant – all without any required input from a human(Fox).

While at first this seems reasonable, it becomes to be an incredibly daunting task as researchers continue to learn more about the structure and operation of the various parts of the human brain. The simple fact that not everything is known about the brain makes attempting to recreate it digitally very difficult. For example, thirty years ago, neuroscience researchers estimated that the human brain contained between three and ten billion neurons. Today, the number is estimated at a hundred billion (Kassan). In addition, “it was assumed that the brain's glial cells, which outnumber neurons by nine times, were purely structural and had no other function. In 2004, it was reported that this wasn't true” (Kassan). To put it simply, modeling the human brain is a difficult, maybe impossible task. Kassan's main point is to pose the question of whether or not we have ever modeled the brain of any organism. No, but to add to it he points out that the brain of a nematode worm (C. Elegans) has been being studied for 40 years. There are several websites and thousands of scientists working on it – a brain of just over three hundred neurons. And even still, it has only been mapped (which, as he points out, is an easier but less useful task than modeling). It has never been modeled – “*and the number of neurons in the human cortex alone is 100 million times larger*...” (Kassan, empahsis his).

In addition, Kassan adds that the program to operate this artificial brain, even if it were possible to create, would be massive. He estimated that even if each neuron in the brain were controlled by only a single line of code, the software would still be 25 million times larger than the biggest software project, Microsoft Windows. He also indicates that as a program increases in size, the probability of error and failure increase as well. In essence, the possibility of the code working is zero.

Kassan’s opinions cast the field of artificial intelligence into a very negative light. It would appear that based on his information and opinions, artificial intelligence will never happen. It may not, at least not in the sense of creating a robot with full, human intelligence, but it is *inevitable* that machines, intelligent at least in a limited scope, will come to play a huge role in our lives. Today, there are already several devices that use some level of artificial intelligence to allow them to perform their tasks better than their purely mechanical ancestors. These ‘smart’ machines, as they’re commonly called, could likely be the sole outcome of the attempts to create artificially intelligent machines. Therefore, it is important to speak of them briefly.

“Smart” devices can serve many purposes, and operate in many different places. Typically, the purpose of these smart devices is to eliminate some of the hard work that people have to do (Graham-Rowe). The tasks they usually do require great precision, or are boring and repetitive so that humans no longer want to do them (humans tend to make more mistakes once they become bored). For example, many factories filled repetitive jobs on assembly lines with robots as they became available. It saved them the money they would have paid employees and robots do not need to take breaks, and they do not make mistakes. In addition, there are a number of robots that are now inside our households. For example, there are robotic vacuum cleaners that can turn themselves on and off according to programmed schedules that will even vacuum around obstacles such as furniture. By including simple computers in devices such as washing machines and dishwashers, these devices are able to complete their tasks better and with even less input from their human owners, saving us time and energy. We watch as cell phones become “smarter”, and are jam-packed with more technology than most of us ever need. It is foreseeable that these types of machines may very likely fulfill even more duties in the future.

Another function of smart devices is to take over jobs that could cause injury or prove fatal to humans (Graham-Rowe). For example, the military uses robots to disarm bombs and to locate land mines. In addition, unmanned aerial vehicles, or UAVs, are miniature spy planes used by the military to scout out positions of enemies or by local agencies to track forest fires. Some UAVs have also been armed with weapons, so that in case they encounter hostility, the remote pilot can return fire (Graham-Rowe). In addition, the robotic soldier, a dream of the Pentagon for 30 years (about $127 billion has been spent on its development) is coming within reach (Vesely). A statement by Johnson, of the robotics effort at the Joint Forces Command Research Center, echoed the message: "Today we have an infantry soldier. We give him a set of instructions; if you find the enemy, this is what you do. We give the soldier enough information to recognize the enemy when he is fired upon. He is autonomous but he has to operate under certain controls. By 2015 we think we can do the same infantry missions with robots" (Vesely). In the case of replacing soldiers with robots, not only is the preservation of human rights an important motivator, but so is money: "Future commitments for soldiers' retirement benefits already total $653bn [billion], a figure the military is unable to come up with if called upon to do so. The current median cost per basic infantry man is $4m [thousand] and growing, the cost of maintaining, or scrapping, a worn-out robot estimated at a tenth of that figure. "To have robots do at least the basic tasks therefore makes economic sense to a US military watching the mounting cost of modern armaments with increasing alarm" (Vesely). NASA has also been creating a robot called the robonaut, which is able to perform dangerous space walks in place of astronauts.

Smart devices are also showing up in hospitals, as more and more doctors prefer to use robots for precision surgery (Graham-Rowe). Examples of these robots includes the DaVinci robotic surgery system - which assists surgeons in keyhole alloowing them greater precision when working on "anything from gall bladder removals and brain surgery to heart bypasses" (Graham-Rowe). Miniature robotic camera capsules can also be used to search inside a patient's body for problems, and some specialized devices can even be used in the cancer screening process.

In addition to smart devices, robots are often used in the research process, and scientists hope that in the future, advanced, intelligent robots will help us come to better understand nature and long-term natural processes like evolution. The beginning of this can be seen today, as computer scientists attempt to gain a better understanding of evolution through advanced computer simulations.

One example of such a simulation was a study conducted at the University of California, Los Angeles (UCLA), which created a simulation of ants that would attempt to find their way down a trail ("Computers Come to Artificial Life"). The "ants" in the experiment, however, were merely pictures on a computer screen. They were not given any precise coding on how to navigate the virtual "trail", but rather were merely assigned random values of how they might attempt it. The end result was that "(w)hen these ants were first let loose on the trail, they did horribly. But some did manage to move a few steps along the trail" ("Computers Come to Artificial Life"). The researchers chose the more successful ones and "mated" them, by splitting their programs in half and recombining them, much like how the process occurs in animals. They would even occasionally change a random value to simulate DNA mutations that sometimes occur. The key to the experiment was that "(b)y the 70th generation, many of the ants could find the end of the trail" ("Computers Come to Artificial Life"). The researchers believe that the ants' programs had "evolved" so that they were able to find the end of the trail.

A second example is actually a program that was written with no intention to be educational at all. Called “Boids”, it is a program that was written by Computer Animator Craig Reynolds. The program itself is a simulation of a pack of birds (“Computers Come to Artificial Life”). In the program, the Boids follow three simple rules on how to fly. First was a “Clumping Rule”, which attempted to keep them flying together as a group. Second is a rule governing their speed, which is supposed to make the boids match the speed of those they are flying near. The third rule tells them not to get too close to each other so they do not collide. When Reynolds programmed these rules into the individual boids and released a large group of them, they acted like real birds birds (“Computers Come to Artificial Life”). In fact, they behaved so much like birds that after releasing his program, ornithologists (people who study birds) began to call him. They had never been able to adequately explain how birds flew together, and they believed that Reynolds may have uncovered some of the underlying principles of bird flight birds (“Computers Come to Artificial Life”).

A third experiment, one that really capitalizes on the capabilities of intelligent machines, was performed by entomologist Laurent Keller, who wanted to examine how something complex like a communication system could arise through evolution (and, if in fact, it did at all) (Coleman). Traditionally, it has been very difficult to study something like this, as social systems are believed to arise over several hundred generations of creatures, and species that are highly sociable, like bees or ants, have long life cycles. As a result, it would take several years to produce results, and unlimited possibilities for the experiment to be compromised. In addition, communication systems leave no evidence that can be gleaned from remains like tooth or skull size. However, Keller was able to create a new, efficient way to get the results needed. Instead of using a creature, he used "s-bots": a specialized type of robot that comes "equipped with wheels for motion, a blue light for signaling, and a "genome" in the form of a computer program" (Coleman). "S-bots are the answer to evolutionary biologists' prayers: they can be programmed to live, undergo selective pressure, and reproduce in two minutes flat" (Coleman). Keller was able to run hundreds of "generations" of the machines through exercises in just a week, and was able to effectively studied how the bots "evolved". In Keller's experiment, the bots were kept in an enclosed environment, along with a source of food and a source of poison. Their goal was to locate the food – they were given points for successfully finding it, and lost points for locating the poison (Coleman). At the end of each "round", the most successful were "breed", much like the ants in the UCLA experiment above – their programs were split, recombined, and occasionally mutated. In the beginning, the bots were able to flash their lights, but were given no instruction as to why. Over the course of the experiment – Keller examined a few things. First, when placed in group scenarios, the bots always evolved a system of communication – they would either flash their lights to indicate the location of food (and attract other bots to their position), or they would learn to signal the location of poison, in which the other bots would stay away. However, since the bots could only flash their lights in one way, they could only choose one system or the other. Even though the location of food signaling is more efficient and advantageous to the group, once one system evolved, it stuck – the bots never switched the system they had evolved once they began using it. It is believed that "(t)his finding helps explain why in the real world, animals sometimes evolve convoluted communication systems when simpler models seem possible" (Coleman). In addition, Keller also performed versions of the experiment in which the bots were not working as a group, but rather were graded on individual performance. In the experiment's biggest twist, the bots learned to "lie" – the bots would use their lights to lure their opponents away from food sources. In this case, the "bots did evolve communication systems, but instead of boosting performance, communication actually dragged down the performance of the group as a whole" (Coleman). It would seem as though "(l)ying is clearly a natural impulse that can arise through evolution, but the robot experiment proved that it's just not as productive as telling the truth" (Coleman).

In addition to being the subject of study for researchers, intelligent robots could one day be the researchers themselves. A "Robot Scientist", created by Ross King of the University of Wales, Aberystwyth, biologist Steve Oliver of the University of Manchester in Britain and six of their colleagues, was designed to test the feasibility of replacing humans with robots when performing "grunt work" science ("Robot Scientist Does Research Without Human Help"). To test their robot, the scientists had it determine the role of a group of genes in the synthesis process of particular amino acids. Though the question had already been solved, it was a perfect of example of what type of work the robot would do if it became a mainstream tool, and it allowed its results to be tested for accuracy. The robot designed a list of all possible hypotheses, and then determined, by analyzing the cost and probability of success, which study it should conduct. It followed through with the experiment, gathered the data, and assessed its results. "Then came the real test. The researchers assigned nine graduate students the same task they had given the Robot Scientist, and told them to perform it as quickly and cheaply as they could. But the robot held its own: there was "no significant difference between the robot and best human performance in terms of the number of [experiments] required" to find the answers. Both the humans and the robot had to conduct about five experiments to determine the function of a gene. Financially, though, the robot came out slightly ahead" ("Robot Scientist Does Research Without Human Help"). The ability of the bot to at least match the graduate students shows that it is possible that the bot, or one like it, could one day find its way into the labs as a common research tool.

Clearly, artificial intelligence is bound to have an impact in our lives in the future. The specialized purposes above easily prove that. The problem with artificial intelligence, however, is the concept of making it all-purpose. This human-like intelligence (or super-intelligence) is where most of the problems with intelligent machines arise. If intelligent humanoid robots are an inevitability (and as shown above, it is far from inevitable), there are several concerns one should think about besides just the clichéd robot rebellion. For example, if these robots did exist, should they be given the same legal rights and responsibilities that are now afforded to humans? In 2004, a mock trial was held to investigate the very idea. In it, the plaintiff was a powerful supercomputer with emotional artificial intelligence that had been used by a fictitious corporation to deal with costumer relations. However, by scanning confidential memos, the computer had learned the company had decided to deactivate it and scrap it to build a new model. The computer had sent an email to some lawyers, asking them to represent it and save it from deactivation (Soskis). The lawyers for the computer, called BINA48, pleaded several statutes under California law (where the trial was held) in order to convince the jury to grant their injunction motion to prevent the company from pulling the plug on BINA48. Among them were laws governing patients on life support and animal cruelty laws. Essentially, they argued that "An entity that is aware of life enough and its rights to protest their dissolution is certainly entitled to the protection of the law" (Soskis). In addition, BINA48 testified at trial, played by a woman who was pretending to be a holographic image generated by the computer. The defense, of course, argued simply that, as a computer, BINA 48 had no right to bring a motion into trial, as machines could only fake emotions, and did not have any real intelligence. The outcome of the "trial" is interesting. "The jury, comprised of audience members, sided overwhelmingly with the plaintiff. But the mock trial judge, played by a local lawyer who is an expert in mental health law, set aside the jury verdict and recommended letting the issue be resolved by the hypothetical legislature. The audience seemed to regard the compromise with some relief, as if their hearts were with BINA48 but their minds with judicial restraint"(Soskis).

Beyond the huge question of whether or not robots should have the same rights as humans, one must also consider the possibility that robots will eventually fill many of the same jobs currently filled by humans. How might that affect the workforce? How will it affect the ability of a person to get a job (especially if they have little or no specialized training)? Some argue that it will not make a huge impact, emphasisng that the goal of robot designers is not to recreate, or rather replace, humans, since the two are fundamentally different in their skills and abilities. They remind us that the invention of the calculator did not eliminate the mathematician, nor did it lessen the importance of math training in schools. What it has done, however, is change what is important, or desirable, in the math career. No longer is the ability to multiply a large number in your head very valued. But the calculator hasn't stolen that ability from us, it has merely freed our mind to pursue smething more worthy – like new formulas or relations – something that a calculator still can't do (Bruemmer).

This is highly arguable though, if you consider the lowly factory line worker who is being replaced by machines all around the world. In addition, an emerging trend is to replace higher level positions in companies with machines as well, such as robotic project managers that can more efficiently locate and allocate resources than their human peers (Brown). In addition, entire industries are being replaced by robots. For example, it is possible to purchase software, TurboTax for example, to help prepare taxes - instead of going to companies to have humans help. As these programs become more powerful and extend into other industries, it will not only be blue collar jobs that robots replace, but white collar jobs as well (Brown). Another emerging trend in business is for companies to try to hire fewer and fewer on-site employees. By reducing or eliminating the need for a physical location of their business, employers can eliminate the need to build and maintain the location, not to mention the cost of insuring the building and the people who work within it (Brown). Consider this: if a job is being done off-site and over the internet, will the companies care if a human is on the other end doing the job as opposed to an intelligent machine posing as a human? (Brown).

One final, though major, concern we should have with the concept of artificially intelligent machines is how to ensure they have an acceptable level of morality. Isaac Asimov, a famous author of science fiction novels including *I, Robot*, laid the groundwork for the thought of artificial morality in his book Runaround. In it, he stated his now-famous Three Rules of Robotics, which were rules that all robots had to follow or else they would stop functioning (their electronic “brains” would short-circuit). The three rules, which AI researchers constantly examine and critique, are as follows:

> *1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
>
> *2. A robot must obey orders given to it by human beings, except where such orders would conflict with the First Law.*
>
> *3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.*

Although these rules are primitive (and of arguable effectiveness), and though they have been modified time and again by authors, movie directors, and even Asimov himself, they represent what many would consider the bare minimum of the morality we would expect from intelligent machines. Likely, individual societies may also want to include additional rules in their robots as their culture and their laws demand. However, if the ability to change what the robot considers moral is left to the individual programmer, it is always possible that these rules could be left out, and that perhaps, the rules could be changed so that a robot would begin to injure people. This is why some AI researchers argue that it is imperative that we create "Friendly AI" – that is, an artificial intelligence that is humane – as soon as possible (Yudkowsky). They argue that as computers become faster, the likelihood of intelligence "emerging" increases. However, although the power of computers is following Moore's Law (computational power doubles every 18 months), our understanding of intelligence has not. As the argument goes, "Moore's Las does make it easier to develop AI without understanding what you're doing, but that's not a good thing" (Yudowsky). The reason it's not a good thing is because although the difficulty in creating AI is decreased, the difficulty in creating *friendly* AI is not. This is because friendly AI does not require more powerful hardware, but rather a better, more complete understanding of the human mind, of emotions, and thought (Yudowsky). The argument is that once hardware has become powerful to run an AI if it was built perfectly, than time is no longer on your side. Because from that point on, hardware will get stronger, allowing for more inefficiencies and errors in the code that will still be able to run fast enough. As hardware gets more powerful, the program can become sloppier, eventually allowing for someone to just guess and get it close enough to work – without an understanding of how it *should* work. If friendly AI cannot be created by that time, “we're in very serious trouble” (Yudowksy).

There are clearly several legitimate concerns about artificially intelligent robots and how they can impact humanity. With all these risks, why would anybody want to continue in this undertaking? How can artificial super-intelligence help us, and is it worth the great risks associated with it? There are several artificial intelligence experts who would answer that question with an emphatic YES. They believe that artificial intelligence will not only be mankind’s greatest invention, it could very well put us on the road to great achievements such as the complete eradication of disease and hunger.

These experts believe that a point in time known as The Singularity is coming upon us, and may very well happen within the current generation’s lifetime. What is The Singularity? Nick Bostrom, director of The Future of Humanity Institute of Oxford University, states it as the moment in time when machine intelligence exceeds that of human intelligence, or to put it another way, becomes super-intelligent. The belief is that this will likely happen several times faster than the time it takes machine intelligence to reach roughly human-level. More broadly, it can be explained by the following excerpt , which is written by Ray Kurzweil, an inventor, scientist, and author, who agrees that The Singularity may not be far off. As he describes it, the Singularity is some point in the future where technology evolves at a rate so fast, we cannot even begin to imagine it today. Technology will have irreversibly altered our exist on the planet. We will combine our brains, our skills, and our technology to create, communicate, and innovate in ways that would appear to us today as magic. He describes this new world as one in which an explosion in machine intelligence and the fields of gene research and nanotechnology has blurred the lines between biological and mechanical, or physical and virtual, so much that they are no longer visible. He believes that these revolutions will allow us to transcend our physical selves, eradicating illness, hunger, and poverty (through the use of nanotechnology as a way to generate anything upon demand). By manipulating our genetic code, we will evolve rapidly, and live for as long as we choose. As he describes it, this, in essence, is The Singularity (Kurzweil).

Kurzweil states that technological progress is on an exponential curve. That is, it starts out so slow as to be almost unnoticeable, but once the bend in the curve is passed, progress becomes explosive and almost unfathomable. He believes that humanity has already passed this bend, and because of this, in the twenty-first century, we will not make the same number of discoveries as we did in the twentieth century – instead, we will make around 1,000 times more. These discoveries, he argues, will be influenced from today's computers and, of course, the inevitable super-intelligent machines. He believes that the technological progress will be split into three major sections: genetics, nanotechnology, and robotics.

From the genetics breakthroughs, Kurzweil argues, we will see, among other things, a greatly expanded use of the 'RNA Interference' technique. This technique manipulates the DNA and RNA in our cells through the use of specially-designed peptides to turn off disease-causing genes and activate those that are beneficial. In essence, we will be able to reprogram our bodies by altering our genes and metabolic processes. How will these techniques impact our lives? According to Nanomedicine researcher Robert Freitas, "eliminating 50% of medically preventable conditions would extend human life expectancy 150 years." Kurzweil continues that "if we were able to prevent 90% of naturally occurring medical problems, we'd live to be more than 1,000 years old.”

The nanotechnology revolution will take shape in the form of nanoscale robots (those whose key components measure 100 nanometers or less in diameter) that are able to be used to further extend our lifespan. Imagine a device implanted into someone’s body that can measure levels of various chemicals and release certain medicines to benignly influence the use of the chemicals in that person’s body. Acting like an artificial pancreas, this device would be far more accurate and less likely to make a mistake. In addition, Kurzweil believes that nanotechnology will also be implanted into our brains, as a way to speed up our slow interneuron pathways to greatly enhance our own brains, and allow us to connect to other brains wirelessly from anywhere in the world. Like a cell phone in your brain, it would be able to connect you to anyone in the world instantly. Or, like an internet connection in your brain, giving you the ability to receive any information at any time from anywhere.

Finally, the robotics revolution will yield the great mechanical super-intelligence that will be, in Kurzweil’s words, “trillions of trillions of times more powerful than unaided human brain power." Once machines become the dominant intelligence on the planet, we will see technological process expand even more quickly, for specific reasons. First, machines are able to share knowledge faster and more efficiently than humans can.

Machines can send digital files instantly, whereas humans share information solely through slow, language-based communication (until we have internet-brains). Additionally, machines are far better suited than humans for solving complex arithmetic problems, such as those in scientific data. As a result, machines will make far better scientists than humans once machine intelligence exceeds that of humans.

It would be easy to write off the beliefs of Kurzweil, Bostrom, and others like them. It's easy because it is hard to imagine a world like the one Kurzweil describes. It sounds like science fiction, or a hastily thought up fantasy world. Naturally, there are several AI researchers (and many in the general public) that do. However, even if their beliefs do not come to fruition, most of us do believe that artificial intelligence is coming, and it will impact us in whatever form it takes. There are risks to be sure, such as the loss of work (or, perhaps, merely a redefinition of what work humans do), the difficulty of assigning rights and legal responsibilities to machines and a way to enforce the associated laws, and science fiction's favorite risk, the possibility of a robot-led rebellion on mankind. There are, however, also potential rewards. The eradication of disease, hunger, and poverty, and the resulting expansion in the human lifespan tops the list, followed by great enhancements in all scientific fields. However, as originally stated, two things should be clear: the goal of creating a machine that has both the intelligence capabilities as well as the emotional capabilities of a human is still far off, and may never happen. It has also been shown how robots that contain specialized intelligence work well and have proven themselves to be great tools in the lab, in the home, and on the battlefield. Why go farther? What reason is there to risk the dangers of granting human-intelligence to a machine? To chase the dreams of immortality, space travel, and the Internet hard-wired into our brains? Most of these things can likely be solved by humans given enough time. Making a machine capable of doing it for us should hopefully, by now, seem foolish and dangerous.

Works Cited

"Artificial Intelligence." Encyclopedia. Reproduced in Today's Science. Facts On File News Services. 27 Sept. 2008 <http://0-www.2facts.com.helin.uri.edu:80>

"Artificial Intelligence Beats Human Genius." Today's Science On File. July 1997. Today's Science. Facts On File News Services. 27 Sept. 2008 <http://0-www.2facts.com.helin.uri.edu:80>

Asimov, Isaac. Runaround. via http://www.rci.rutgers.edu/~cfs/472_html/Intro/NYT_Intro/History/Runaround.html. Accessed 10/25/09.

Bauer, Peter; Stephan Nouak; Roman Winkler. "A Brief Course in Fuzzy Logic and Fuzzy Control". 4 Dec, 1996. Accessed 31 Jan, 2009. http://www.esru.strath.ac.uk/Reference/concepts/fuzzy/fuzzy.htm

Bostrom, Nick. "Creating Superintelligence Involves Less Risk than Waiting." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Brown, Arnold. "The Impact of Robots on Employment." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Bruemmer, David. "Humanoid Robots Cannot Replace Humans." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Coleman, Elisheva. "From the Mouths of Bots: Evolving Robots Learn to "Talk"." Today's Science. Mar. 2007. Facts On File News Services. 27 Sept. 2008 <http://0-www.2facts.com.helin.uri.edu:80>

"Computers Come to Artificial Life." Today's Science On File. Mar. 1993. Today's Science. Facts On File News Services. 27 Sept. 2008 <http://0-www.2facts.com.helin.uri.edu:80>

Fox, Douglas. "Building Intelligent Machines by Copying Living Brains." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia]Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Graham-Rowe, Duncan. "The Variety of Present-Day Robots." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 28 Sep. 2008

Kassan, Peter. "Duplicating Human Intelligence Is a Mirage." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing

Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Kurzweil, Ray. "Future Technology Will Benignly Alter Human Existence." Opposing Viewpoints: Technology and Society. Ed. David Haugen and Susan Musser. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Malcolm, Chris. "Technology Will Not Threaten Society." Opposing Viewpoints: Techology and Society. Ed. Auriana Ojeda. San Diego: Greenhaven Press, 2002. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Manes, Stephen. "Your $1000 Computer May Not Be So Grand." Http://www.nytimes.com. 28 Jan. 1997. New York Times. 19 Jan. 2009. http://query.nytimes.com/gst/fullpage.html?res=9B0DE7D8103AF93BA15752C0A961958260

""Robot Scientist" Does Research Without Human Help." Today's Science On File. Mar. 2004. Today's Science. Facts On File News Services. 27 Sept. 2008 <http://0-www.2facts.com.helin.uri.edu:80>

Soskis, Benjamin. "Should Intelligent Machines Have Legal Rights?." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Schaeffer, Jonathan; Neil Burch; Yngvi Bjornsson; Akihiro Kishimoto; Martin Muller; Robert Lake; Paul Lu; Steve Sutphen. "The checkers terminator." The Wilson Quarterly. 32. 1 (Wntr 2008): 81(2). Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 27 Sep. 2008

Vesely, Milan. "Robots for Fighting Future Wars." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008

Yudkowsky, Eliezer. "Superintelligence Without Supermorality Would Be Dangerous." Contemporary Issues Companion: Artificial Intelligence. Ed. Sylvia Engdahl. Detroit: Greenhaven Press, 2007. Opposing Viewpoints Resource Center. Gale. Johnson and Wales University. 14 Sep. 2008