If machines are capable of doing almost any work humans can do, what will humans do?

April 15, 2013

Moshe Y. Vardi, George Professor in Computational Engineering, Director of the Kennedy Institute for Information Technology, Rice University

Abstract: *Over the past 15 years Artificial Intelligence has made a remarkable progress. In 1997, IBM's Deep Blue program beating world champion Gary Kasparov. In 2005, a Stanford autonomous vehicle won a DARPA Grand Challenge by driving 131 miles along an unrehearsed desert trail. Two years later a CMU autonomous vehicle won a DARPA Urban Challenge by driving 55 miles in an urban environment while avoiding traffic hazards and obeying traffic laws. And then, in 2011, IBM's Watson program defeated decisively the two greatest in Jeopardy! While AI has been proven to be much more difficult than believed by its early pioneers, its inexorable progress over the past 50 years suggests that Herbert Simon was probably right when he wrote in 1956 "machines will be capable ... of doing any work a man can do." I do not expect this to happen in the very near future, but I do believe that by 2045 machines will be able to do if not any work that humans can do, then, at least, a very significant fraction of the work that humans can do. The following question, therefore, seems to be of paramount importance. If machines are capable of doing almost any work humans can do, what will humans do?*

A major debate among economists has been flaring over the past year or so regarding the impact of automation on jobs. The traditional approach by economists is highly skeptical of Luddism, which is defined as distrust or fear of the inevitable changes brought about by new technology. Such a position was expressed by Kenneth Rogoff, of Harvard University, who wrote, “Since the dawn of the industrial age, a recurrent fear has been that technological change will spawn mass unemployment. Neoclassical economists predicted that this would not happen, because people would find other jobs, albeit possibly after a long period of painful adjustment. By and large, that prediction has proven to be correct.” But in a December 2012 working paper for the National Bureau of Economic Research, called “Smart Machines and Long-Term Misery”, Jefferey Sachs, of Columbia University, and Laurence Kotlikoff, of Boston University, posed the question, "What if machines are getting so smart, thanks to their microprocessor brains, that they no longer need unskilled labor to operate?" After all, they point out, "Smart machines now collect our highway tolls, check us out at stores, take our blood pressure, massage our backs, give us directions, answer our phones, print our documents, transmit our messages, rock our babies, read our books, turn on our lights, shine our shoes, guard our homes, fly our planes, write our wills, teach our children, kill our enemies, and the list goes on.” Lawrence H. Summers, of Harvard University, said recently, “As economists like to explain, the system will equilibrate at full employment, but maybe the way it will equilibrate at full employment is there’ll be specialists at cleaning the shallow end and the deep end of rich people’s swimming pools. And that’s a problematic way for society to function.”

There is considerable evidence that the neo-Luddites may be justified in their concerns. Eric Brynjolfsson and Andrew McAfee, of MIT, recently wrote: "For several decades after World War

Il the economic statistics we care most about all rose together here in America as if they were tightly coupled. G.D.P. grew, and so did productivity — our ability to get more output from each worker. At the same time, we created millions of jobs, and many of these were the kinds of jobs that allowed the average American worker, who didn't (and still doesn't) have a college degree, to enjoy a high and rising standard of living. Productivity growth slowed in the 1970s but revved up again in the 1990s and has stayed strong most years since. But … productivity growth and employment growth started to become decoupled from each other at the end of that decade. We are creating jobs, but not enough of them. The employment-to-population ratio, or percentage of working-age people that have work, dropped over 5 points during the Great Recession, and has improved only half a point in the three and a half years since it ended."

While the neo-Luddite economists marshal their data and argue that "this time it is different," and the neoclassical economists argue that "this time it is not different", it may be useful to consider this debate from the perspective of computational intelligence, going back to the foundations of the field.

Alan Turing first addressed the issue of computational intelligence in an unpublished 1948 paper titled "Intelligent Machinery", where he argued that machines can achieve intelligence. He returned to this topic in 1950 in what is probably his most well-known paper, "Computing Machinery and Intelligence," where he proposed the "Imitation Game" as an operational definition for machine intelligence. In this game, now known as the "Turing Test", a computer has to engage in an online conversation with a human and convince the human that it, the computer, is also human.

Turing was not the first to think about computational intelligence, more popularly referred today by the term "artificial intelligence" (AI), coined by John McCarthy in 1955. William Stanley Jeavons, the British logician who built a mechanical logic machine, known as the "Logic Piano", in 1869, wrote about his machine: "The machine represents a mind endowed with powers of thought, but wholly devoid of knowledge. ... It cannot be asserted indeed that the machine entirely supersedes the agency of conscious thought." The American philosopher Charles S. Peirce wrote in 1887: "Precisely how much the business of thinking a machine could possibly be made to perform, and what part of it must be left to the living mind is a question not without conceivable practical importance." Nevertheless, Turing's 1950 paper is indeed the first deep philosophical investigation of the possibility of artificial intelligence. While the Turing Test has been rather under-influential in the history of AI, Turing does deserve the credit for putting the question of general machine intelligence so squarely on the table.

The main focus of Turing's 1950 paper is actually not on the Imitation Game but on the possibility of machine intelligence. Turing carefully analyzed and rebutted arguments against machine intelligence. He also stated his belief that we will see machine intelligence by the end of the 20th Century, writing "I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted." We now know that he was too optimistic. We seem quite far from describing today's computers as thinking machines. Turing was not the only one to be too optimistic. Many of the early AI pioneers were brimming with unbounded optimism. Allen Newell wrote in 1958 that "within ten years a digital computer will be the world's chess champion," and Marvin Minsky wrote in 1967 that "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved." Perhaps because of such over-optimism, AI has suffered from repeated "AI Winters", periods that were characterized by slow progress and death of research funding. AI researchers refer to the "First AI Winter", 1974-1980, and "Second AI Winter", 1987-1993.

In the late 1990s, however, AI seems to have turned a corner. 1997 saw IBM's Deep Blue program beating world champion Gary Kasparov 3 1/2 - 2 1/2 in what was surely the most dramatic chess match of the 20th century. In 2005, a Stanford autonomous vehicle won a DARPA Grand Challenge by driving 131 miles along an unrehearsed desert trail. Two years later a CMU autonomous vehicle won a DARPA Urban Challenge by driving 55 miles in an urban environment while adhering to traffic hazards and traffic laws. And then, in early 2011, IBM's Watson program decisively defeated the two greatest Jeopardy! champions, Brad Rutterand and Ken Jennings. Perhaps Theodore Kaczynski, the infamous Unabomber, was right when he wrote in 1995: "Let us postulate that the computer scientists succeed in developing intelligent machines that can do all things better than human beings can do them."

In fact, Turing's 1950's philosophical analysis is still compelling today. I see no reason to believe that machines cannot be as intelligent as humans. In fact, if we reject metaphysical arguments, the human brain is an intelligent machine, albeit a biological one, and organized quite differently than modern computers. Furthermore, even if Roger Penrose is right in his argument that the brain employs quantum computation, there is no a priori argument why we would not be able to build an analogous machine.

The question of what happens when machines get to be as intelligent as and even more intelligent than people seems to occupy mostly science-fiction writers. The Terminator movie trilogy, for example, 1984-2003, featured Skynet, a self-aware artificial intelligence that served as the trilogy's main villain, battling humanity through its Terminator cyborgs. Among technologists, it is mostly "Singularitarians" who think about the day when machine will surpass

humans in intelligence. The term "singularity" as a description for a phenomenon of technological acceleration leading to "machine-intelligence explosion" was coined by the mathematician Stanislaw Ulam in 1958, when he wrote of a conversation with John von Neumann concerning the "ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue." More recently, the concept has been popularized by the futurist Ray Kurzweil, who pinpointed 2045 as the year of singularity. Kurzweil has also founded Singularity University and the annual Singularity Summit.

It is fair to say, I believe, that Singularitarians are not quite in the mainstream. Perhaps it is due to their belief that by 2045 humans will also become immortal and be able to download their consciousness to computers. It was, therefore, quite surprising when in 2000 Bill Joy, a very mainstream technologist as co-founder of Sun Microsystems, wrote a "heretic" article entitled "Why the Future Doesn't Need Us" for Wired magazine. "Our most powerful 21st-century technologies--robotics, genetic engineering, and nanotech--are threatening to make humans an endangered species," he wrote. Joy's article was widely noted when it appeared, but it seemed to have made little impact.

It is in the context of the Great Recession that people started noticing that while machines have yet to exceed humans in intelligence, they are getting intelligent enough to have a major impact on the job market. In his 2009 book, “The Lights in the Tunnel: Automation, Accelerating Technology and the Economy of the Future”, Martin Ford wrote “The current crisis has been perceived as primarily financial in origin, but is it possible that ever advancing technology is an unseen force that has contributed significantly to the severity of the downturn? More importantly, what economic impact will technological acceleration have as we anticipate recovery from the current crisis--and in the years and decades ahead? What will the economy of the future look like?.”” In their 2011 book, “Race Against The Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy”, authors Brynjolfsson and McAfee, argued that "technological progress is accelerating innovation even as it leaves many types of workers behind."

Indeed, over the past 30 years, as we saw the personal computer morph into tablets, smartphones, and cloud computing, we also saw income inequality grow worldwide. While the loss of millions of jobs over the past few years has been attributed to the Great Recession, whose end is not yet in sight, it now seems that technology-driven productivity growth is at least a major factor. Such concerns have gone mainstream in the past year, with articles in newspapers and magazines carrying titles such as "More Jobs Predicted for Machines, Not People," "Marathon Machine: unskilled workers are struggling to Keep up with technological

change," "It's a Man vs. Machine Recovery," and "The Robots Are Winning", with even prominent economists such as Paul Krugman writing about “The Rise of the Robots”.

While AI has been proven to be much more difficult than Turing and other early pioneers believed, its inexorable progress over the past 50 years suggests that Herbert Simon was probably right when he wrote in 1956 "machines will be capable ... of doing any work a man can do." I do not expect this to happen in the very near future, but I do believe that by 2045 machines will be able to do if not any work that humans can do, then a very significant fraction of the work that humans can do. Bill Joy's question deserves therefore not to be ignored: does the future need us? By this I mean to ask, if machines are capable of doing almost any work humans can do, what will humans do? As I have been raising this question over the past year, I have been getting various answers to this question, but I find none satisfying.

Simon himself addressed his question in a 1965 book, “The Shape of Automation for Men and Management”. Simon described himself as a “technical radical”, believing that computers will be able to do anything men can do, and “economic conservative”, seeing computer automation as simply the continuation of the Industrial Revolution, with similar consequences. His economic analysis, however, is not as compelling today as it may have been in 1965. He wrote then that “Empirical case studies of automation do not reveal any general tendency towards either the upgrading or the downgrading of job skill requirements.” Today’s labor economists are much less sanguine. As we saw, Sacks and Kotlikoff argueed that computer automation is pushing for up-skilling of the job market. Furthermore, Simons assumed that increased productivity and per capita income are coupled, which use to be the case in 1965. But Brynjolfsson and McAfee documented that, starting in the early 1980s, productivity growth and per capita incomes started decoupling.

Simon believed that technology has been destroying jobs since the start of the Industrial Revolution, yet new jobs are continually created. This has been happening for about 200 years and there is no reason to believe that “This time it is different.” The AI Revolution, however, is different, I believe, than the Industrial Revolution. In the 19th Century machines competed with human brawn. Now machines are competing with human brain. Robots combine brain and brawn. We are facing the prospect of being completely out-competed by our own creations. As an example, less than 10 years after autonomous vehicles won the DARPA Grand Challenge and Urban Challenge, several states in the US are rewriting their laws to allow for automated cars. Millions of driving jobs are going to be eliminated in the next couple of decades. I suspect that in 25 years driving by people will simply look quaint. An Associated Press analysis of employment data from 20 countries found that millions of mid-skill, mid-pay jobs already have disappeared over the past five years, and they are the jobs that form the backbone of the middle class in developed countries. “Technological change is more encompassing and moving faster and making it harder and harder to find things that people have a comparative advantage

in versus machines, said David Autor, of MIT, who has studied the loss of mid-pay jobs to technology. If one accepts the proposition that computational intelligence is inexorably gaining on human intelligence and that robots are making inexorable progress in their ability to operate in the physical world, and it becomes harder and harder not to accept this proposition, then it gets harder and harder to imagine what jobs will be left for humans to do.

A counter-argument is that the prevalence of low-cost industrial robots would lead to a *rise* in employment. Such robots can lower the cost of manufacturing; bringing back to the US industries that have left due to offshoring, or keeping in the US industries that otherwise would have left. But this argument is only about comparative advantage between countries. In the aggregate jobs are still lost. If manufacturing that was offshored to China returns to the US, then jobs are lost in China. Overall, low-cost robots may increase productivity without creating jobs.

Another typical answer to the job-loss argument is that if machines will do all of our work, then we will be free to pursue leisure activities. The economist John Maynard Keynes addressed this issue already in 1930, when he wrote "The increase of technical efficiency has been taking place faster than we can deal with the problem of labour absorption." Keynes imagined 2030 as a time in which most people worked only 15 hours a week, and would occupy themselves mostly with leisure activities. I do not find this to be a promising future. First, if machines can do almost all of our work, then it is not clear that even 15 weekly hours of work will be required. Second, while wealthy people may be able to enjoy satisfying, Downton Abbeyesque leisure-filled life, the same may not applied to those who live off unemployment insurance. Furthermore, I do not find the prospect of leisure-only life appealing; I believe that work is essential to human well-being. Third, our economic system would have to undergo a radical restructuring to enable billions of people to live lives of leisure. The unemployment rate in the US is currently just below 8% and is considered to be a huge problem. What will happen to our society if the unemployment rates climbs to 25% or 50%? The balance in our economy between labor and capital is already increasingly shifting further from labor towards capital. One can imagine a future society of a small number of haves and a large number of have-nots, supported, say, by government subsidies. This is reminiscent of "*panem et circenses",* the Roman practice of free bread and entertainment to the masses.

Rodney Brooks, of MIT, has argued that given the large number of aging elders around the world, we will need a major improvement in elder-care productivity in years to come. Robots will provide the only viable solution to this impending elder-care crisis, he argues. Robots may indeed become the standard elder-care givers in years to come, but the picture of a combination of unemployed young and middle-age people and robot-cared elders seems dystopic to me,

Some people argue that these concerns about computational intelligence jobs apply only to a future that is so far away that we need not worry about it. I find this answer to be unacceptable. 2045 is merely a generation away from us. Global-warming concerns are also about the year 2050. We cannot shirk responsibility from concerns for the welfare of the next generation. Of course, there is another possibilities, raised by some Singularitarians, which is that by 2045 will merge and there will be no more “us and them”. So our future is to be assimilated into the “Borg collective”, in Star Trek terminology—another prospect that I find dystopic .

There have been some recent calls, for example, by Clayton Christensen, for rethinking the foundations of our market economy, focusing on innovation that creates jobs rather than in innovation that only destroys jobs. This is an admirable sentiment, but we cannot focus on innovation *for* jobs, unless we understand what jobs will be available to humans when machines can do most of the jobs currently done by humans. It seems intuitive that jobs that require intense interpersonal interaction will be resistant to automation and “robotification”, for example, caring for young ones. But if Brooks can build robots to take care of elders, I assume he can also build robots to take care of young ones. Isaac Asimov’s nanny-robot Robbie comes to mind.

In 2000, Bill Joy advocated a policy of *relinquishment*: "to limit development of the technologies that are too dangerous, by limiting our pursuit of certain kinds of knowledge." I am not ready to go that far, and I doubt that such an approach is practical, but I believe that just because technology can do good, it does not mean that more technology is always better. Turing was what we call today a "techno-enthusiast", writing in 1950 that "We may hope that machines will eventually compete with men in all purely intellectual fields ... we can see plenty there that needs to be done." But his incisive analysis about the *possibility* of machine intelligence was not accompanied by an analysis of the *consequences* of machine intelligences. It is time, I believe, to put the question of these consequences squarely on the table. We cannot blindly pursue the goal of artificial intelligence without pondering its consequences. If machines are capable of doing almost any work humans can do, what will humans do?

I believe that this question of the future of humanity in the face of intelligent machinery is one of the central challenges that are now facing humanity, akin perhaps to the climate-change challenge. In both challenges the consequences of technology threaten the welfare of humanity. The climate-change debate shows how hard it is to deal with such challenges. The severity of the challenge on one hand, and the degree to which we attached to the technology make it exceedingly difficult not only to mount a response to such challenges but even to reach consensus on their very existence.

Bill Joy called for relinquishment of technology. If relinquishment is impractical, what should be the public-policy response to the artificial-intelligence challenge? Most fundamentally, we need

to develop a deeper understanding about the impact of automation and information technology on labor. Over the past 30 years our economy has witnessed stagnating median income, growing income disparity, declining labor to capital ratio, and the like. To what extent are these trends driven by technology, rather than by macro-economic forces such as globalization and tax policy? While globalization is with us to stay, it may be hard, but not impossible to change tax policy. We need to re-examine our tax policies and question their current positive bias towards capital investment rather than job creation.

Most importantly, we must heed Brynjolfsson and McAfee's warning that, "There is no economic law that says that everyone, or even most people, automatically benefit from technological progress." It is fair to assume that technological progress is unstoppable. What we make of it is up to us.