Moshe Y. Vardi

The Moral Imperative of Artificial Intelligence

The big news on March 12 of this year was of the Go-playing AI-system AlphaGo securing victory against 18-time world champion Lee Se-dol by winning the third straight game

of a five-game match in Seoul, Korea. (AlphaGo is a program developed by DeepMind, a British AI company acquired by Google two years ago.) After Deep Blue's victory against chess world champion Gary Kasparov in 1997, the game of Go was the next grand challenge for game-playing artificial intelligence. Go has defied the brute-force methods in game-tree search that worked so successfully in chess. In 2012, *Communications* published a Research Highlight article by Sylvain Gelly et al. on computer Go, which reported that "Programs based on Monte-Carlo tree search now play at human-master levels and are beginning to challenge top professional players." AlphaGo combines tree-search techniques with search-space reduction techniques that use deep learning. Its victory is a stunning achievement and another milestone in the inexorable march of AI research.

By using deep–learning techniques to prune the search tree, AlphaGo can be said to augment brute-force search with "intuition," which it has developed by playing numerous games against itself. (Google said AlphaGo does not use Lee's games as its training data.) By relying on learned "intuition," AlphaGo is able to overcome the so-called *Polanyi's Paradox*. The philosopher Michael Polanyi observed in 1966, "We can know more than we can tell ... The skill of a driver cannot be replaced by a thorough schooling in the theory of the mo-

torcar." Some labor economists have viewed Polanyi's Paradox as a major barrier for AI, arguing it implies a limit on its potential to automate human jobs. AlphaGo's victory demonstrates machine learning provides a path around that barrier.

Indeed, the automation of driving has been a major challenge for AI research over the past decade. In 2004, economists argued driving was unlikely to be automated in the near future due to Polanyi's Paradox. A year later, a Stanford autonomous vehicle won a DARPA Grand Challenge by driving over 100 miles along an unrehearsed desert trail. Now, more than a decade later, both technology companies and car companies vigorously pursue the automation of driving. I expect the technical challenges to be resolved in the coming decade.

In 1843, Ada Lovelace, well known for her work on Charles Babbage's Analytical Engine, wrote to Babbage that she wished to see computing technology developed for "the most effective use of mankind." It is difficult for me to think of any computing technology other than automated driving that can be deployed in a decade or two and with such benefit for humanity. About 1.25 million people worldwide die from car accidents every year. Over 90% of these accidents are caused by human error. By automating driving, we could save over a million lives a year, as well as avoid countless injuries.

At the same time, the automation of driving would have a huge disruptive effect on the global economy. Existing industries will shrivel, and whole new industries will rise. In the U.S., close to 10% of all jobs involve operating a vehicle and we can expect to see the majority of these jobs disappear. The human cost of such a profound change cannot be underestimated. As a precedent we can see what happened to U.S. manufacturing over the past 35 years. While manufacturing output in constant dollars is at an all-time high, manufacturing employment peaked around 1980 and is today lower than it was in 1947. The disappearance of millions of jobs due to automation may explain a recent, rather shocking, finding by economists Angus Deaton, winner of the 2015 Nobel Memorial Prize in Economic Science, and Anne Case, that mortality for white middle-aged Americans has been increasing over the past 25 years, due to an epidemic of suicides and afflictions stemming from substance abuse.

Thus, the automation of driving would be hugely beneficial, saving lives and preventing injuries on a massive scale. At the same time, it would have a profoundly adverse impact on the labor market. In the balance, life saving and injury prevention must take precedence, and we have a moral imperative to develop and deploy automated driving. The solution to the labor problem will not be technical, but sociopolitical. As computing professionals, we also have a moral imperative to acknowledge the adverse societal consequences of the technology we develop and to engage with social scientists to find ways to address these consequences.

Follow me on Facebook, Google+, and Twitter.

Moshe Y. Vardi, EDITOR-IN-CHIEF