Is Research in Intelligence an Existential Risk?

By Tomaso Poggio

December 15, 2014

Background

Recent months have seen an increasingly public debate taking form around the risks of AI (Artificial Intelligence). A letter signed by Nobel prizes and other physicists defined AI as the top existential risk to mankind. More recently, Tesla CEO Elon Musk has been quoted saying that it is "potentially more dangerous than nukes." Physicist Stephen Hawking told the BBC that "the development of full artificial intelligence could spell the end of the human race". And of course recent films such as *Her* and *Transcendence* have reinforced the message. Thoughtful comments by experts in the field such as Rod Brooks, Oren Etsioni and others have done little to settle the debate.

As the Director of a new multi-institution, NSF-funded and MIT-based Science and Technology Center — called the Center for Brains, Minds and Machines (CBMM) — I am arguing here on behalf of my collaborators and many colleagues, that the terms of the debate should be *fundamentally rephrased*. Our vision of the Center's research integrates cognitive science, neuroscience, computer science, and artificial intelligence. Our belief is that understanding intelligence and replicating it in machines, goes hand in hand with understanding how the brain and the mind perform intelligent computations. The convergence and recent progress in technology, mathematics, and neuroscience has created a new opportunity for synergy across fields. The dream of understanding intelligence is an old one. Yet, as the debate around AI shows, now is an exciting time to pursue this vision. Our mission at CBMM is thus to establish an emerging field, the Science and Engineering of Intelligence. This integrated effort should ultimately make fundamental progress with great value to science, technology, and society. We believe that we must push ahead with research, not pull back.

A top priority for society

The problem of intelligence — what it is, how the human brain generates it and how to replicate it in machines — is one of the great problems in science and technology, together with the problem of the origin of the universe and of the nature of space and time. It may be the greatest of all because it is the one with a large *multiplier effect* — almost any progress on making ourselves smarter or developing machines that help us think better, will lead to advances in all other great problems of science and technology. Research on intelligence will eventually revolutionize education and learning. Systems that recognize how culture influences thinking could help avoid social conflict. The work of scientists and engineers could be amplified to help solve the world's most pressing technical problems. Mental health could be understood on a deeper level to find better ways to intervene. In summary, research on intelligence will help us understand the human mind and brain, build more intelligent machines, and improve the mechanisms for collective decisions. These advances will be critical to future prosperity, education, health, and security of our society. This is the time to greatly expand research on intelligence, not the time to withdraw from it.

The Science of Intelligence

We are often misled by “big”, somewhat ill-defined, long used words. Nobody can give a precise, abstract definition of what intelligence is. Was Eniac intelligent because it could perform numerical calculations much faster than humans? Is a fly intelligent because its flying acrobatics in chasing potential mates is better than any human pilot? Is a colony of ants intelligent?

Turing’s intuition was in the right direction. With his *Turing test* he tried to provide an operational definition of a specific form of intelligence — human intelligence. Even so, there are many facets of human intelligence — many different competences, many different problems. The word *intelligence* can be misleading in this context, like the word life was during the first half of the last century when popular scientific journals routinely wrote about the problem of life, as if there was a quintessential aether-like substratum of life waiting to be discovered to unveil the mystery. Of course, speaking today about the problem of life sounds amusing: biology is a science dealing with many different great problems, not just one.

I believe the same is true for *human intelligence*: it is not one but many problems, not one but many Nobel prizes. Marvin Minsky referred to this state of affairs using the term *Society of Minds*. A whole science is required to make progress on solving them and to develop the related technologies of intelligence.

For this reason, CBMM has formulated a open-ended set of *Turing*$^{++}$ *questions* in order to measure scientific progress in the field. The plural in questions is to emphasize

that there are many different intelligent abilities that have to be characterized, and possibly replicated in a machine, from basic visual recognition of objects, to the identification of faces, to gauge emotions, to social intelligence, to language and much more. The term *Turing*$^{++}$ is to emphasize that a quantitative model must match human behavior and human physiology — the mind and the brain. The requirements are thus well beyond the original *Turing test*.

Not afraid but watchful

Since intelligence is a whole set of great problems, it is highly unlikely we will see breakthroughs on each and all of them at the same time by a single genius. There is therefore little reason to fear the sudden appearance of a super-human and ultimately anti-human AI as described in *Transcendence*. However, the technologies that are emerging and will emerge over time from the science and the technology of intelligence will be many and they will be powerful and therefore potentially dangerous in their use and misuse, as all technologies are. We need to be responsible and watchful.