

A Science of Intelligence

By [Christof Koch](#), [Tomaso Poggio](#)

July 5, 2015



We are in the midst of a revolution in machine intelligence, the engineering of getting computers to perform tasks that, until recently, could only be done by people. You can speak to your smart phone and it answers back, software identifies faces at border-crossings and labels people and objects in pictures posted to social media. Algorithms can teach themselves to play Atari video games. A camera and chip embedded into the front view-mirror of top-of-the-line sedans let the vehicle drive autonomously on the open

road.

The revolution has been driven by machine learning techniques developed during the last two decades, by Moore's law and by continual improvements in machine architecture. Deep learning and related learning techniques are all the rage and both the industry behemoths as well as tiny start-ups madly scramble to recruit computer science majors with a background in these engineering skills.

Progress seems so dizzying that a few far-sighted futurists, technologists and philosophers are worrying about where this swift ascension of machines is taking us. Will artificial intelligence evolve to be cleverer than any human, becoming fully autonomous in the process, thereby escaping our control and ultimately threatening the very existence of our species? After all, the reason humans dominate the natural world is not because we are faster, or stronger, let alone wiser than other animals, but because we are smarter. By inventing ever more intelligent machines, are we walking into the whirling blades? Will this be our final invention?

How realistic are these Jeremiahs? To assess our current situation, it is natural to ask what is meant by "intelligence" (from the Latin verb *intelligere*, to comprehend or perceive). Unfortunately, intelligence has no universal agreed upon definition.¹

We call a "person" intelligent if they possess common sense, can flexibly adapt to new situations (e.g., a new work environment), can learn new skills, faces or knowledge, and can deal effectively with uncertainty while planning (e.g. planning for college or for retirement). What about a single-cell microbe, such as *E. coli*, which shows limited intelligence by moving toward a sugar-rich solution? Is this intelligence? Is the 1950s Eniac, one of the first computers which could do numerical integrals much faster than

any mathematician, actually intelligent? The truth is that the question of *What is intelligence* is not well-defined as there are too many different answers, an infinite numbers of different kinds of intelligence. Instead, let us focus on a question that is well defined and can be posed in a meaningful way. Let us focus on *What is human intelligence*.

Enter the Turing Test. Invented by the mathematician Alan Turing in 1950 to answer the question “Can machines think?”, it replaces this metaphysically loaded and ill-defined question with a pragmatic imitation game. An agent can be said to have passed the test if it can’t be distinguished from a human being. The Turing test has been popularized in such movies as *Blade Runner* and *Ex Machina*.

Consider the domain of vision and visual intelligence (Figure 1). A contemporary machine vision system could identify all of its faces, even matching them to specific identities, and many of the objects in the scene. However, it would flounder if confronted by an open-ended variety of questions that most children could answer.

What is happening in the scene? What are the various actors doing? Are they friends, enemies or strangers? Why are they there? What will happen next? In which season was this picture taken? In approximately what decade? Where was this picture taken? Is this a real image or a fake one?

What we mean by “understanding an image” entails that we are, at least in principle, able to answer these questions not just with yes and no but with an explanation. Thus we need to develop algorithms that provide explanations of the type humans do. But we really need to achieve more than just simulate the brain’s output, more than what Turing asked. We need to understand what understanding an image by a human brain means. We need to understand the algorithms used by the brain. But we also need to understand the circuits that run these algorithms if we want to be sure that our machine is not just faking the output of a human brain by using a giant look-up table of what people usually do in similar situations, as hinted at the end of the movie *Ex Machina*. Understanding a computer means understanding the level of the software *and* the level of the hardware. Scientific understanding of human intelligence requires something similar — understanding of the mind *as well as* of the brain.

Because of all the different levels of understanding that are involved in answering question on images, we call these *Turing ++ Questions* (also refer to Koch & Tononi, IEEE Spectrum 2008 who use very similar definition for understanding an image). We need to retain the operational aspects of the Turing Test but add to it by moving beyond the level of the behavior to include the level of the underlying neuroscience. What are the internal constraints of an intelligent mind? What role

does learning play? How is learning implemented in neurons and synapses in the brain? What about motivations? What are the ontogenetic and phylogenetic origins of human intelligence? What about the social components of intelligence and which parts of the brain are responsible for them? In particular, which areas of the cortical sheet, a 2+epsilon-dimensional computational tissue, highly developed in humans, are responsible for which cognitive components underpinning intelligence? Is there a single algorithm running in different parts of cortex?

While we go on designing AIs, let us also understand our own intelligence. Ultimately this means answering mysteries that are as old as mankind:

- how does intelligence develop in young children?
- what are the brain regions and specific circuits underlying the various cognitive capabilities involved in answering the Turing ++ questions?
- can we build algorithms that can answer these questions for an arbitrary image or video?
- what are the algorithms and the circuits underlying social intelligence?
- What is the relationship between intelligence and consciousness (here the ability to experience anything, such as the sight of an image)? While they are closely allied in humans, it is conceivable that we can build a truly human-level machine intelligence without any sentience, a zombie.

In summary, we argue that the current focus on the engineering of intelligence, at its heart a problem in the design of better learning algorithms, must be complemented by scientific investigations of the only extant form of natural intelligence – that of *homo sapiens*. It is only once we better understand the biology of our own intelligence that we will be in a much better situation to understand the promises and pitfalls of artificial intelligence. In short, the *engineering of intelligence*, which is now attracting great amounts of resources and talent, calls for an equally grand and synergetic effort — on the *science of human intelligence*.

¹ Legg & Hutter (2006) list 71 definition from philosophy, psychology and computer science for ‘intelligence’.