
World Trends & Forecasts

Artificial
Intelligence

Biodiversity

Families

Criminal Justice

Values

Artificial Intelligence | Sci/TECH

Are You Smarter Than a Sixth-Generation Computer?

Tests for measuring nonhuman intelligence are needed in order to track development.

By Richard Yonck

Despite the amazing achievement of supercomputers such as IBM's *Jeopardy* champion DeepQA (aka Watson), we do not yet call them "intelligent." But what about the next generation of supercomputers, or the ones that come after that?

Before we can make forecasts about machine intelligence, we will need a gauge beyond simple metrics such as petaflops and data-transfer rates. We need to establish a standard metric of machine intelligence.

The idea of testing artificial intelligence goes back to Alan Turing and the eponymously named Turing test. Essentially, the Turing test involved engaging unseen human and machine participants in a text-based conversation. If the judges were unable to correctly identify the AI, then it would be said to have passed. Pass or fail: an all or nothing result.

Unfortunately, while this test is potentially useful for determining *human equivalence*, it's generally agreed that this isn't the only form of intelligence. A dolphin or a chimpanzee could never pass such a test, and yet both exhibit considerable intelligence. It's just that the nature and level of their intellects differ from that of humans.

The same could be said of machine intelligence. Just because human equivalence hasn't yet been achieved *in silico*, doesn't mean that rudiments of intelligence don't exist. Additionally, decades from now, an

artificial general intelligence, or AGI, may be too dissimilar from the human mind to pass the Turing test, even though it might be very superior to us in most other ways.

For more than a century, psychometric tests have existed for people. While some may argue about the merits of assigning a numerical value to the intelligence of individuals, the fact remains that these tests have resulted in considerable knowledge about the distribution of intelligence in our species. Of course, such tests can't be applied to nonhumans. So how does one develop a test suitable for machines?

Over the years, a number of tests of machine intelligence have been proposed. Several, such as Linguistic Complexity and Psychometric AI, suffer from the same shortcomings as the Turing test, in that they test for human equivalence. Many other theorized tests aren't mathematically rigorous enough. To test a nonhuman and rate it on any sort of meaningful scale, we must accurately assess the complexity of the question or challenge set before it.

Recently, a framework for creating mathematically rigorous challenges has been conceived. This is described in a paper titled "Measuring Universal Intelligence: Toward An Anytime Intelligence Test" by José Hernández-Orallo and David L. Dowe, published in *Artificial Intelligence*. This test of "universal intelligence" is grounded in algorithmic information theory and complexity theory in order to

structure its challenges. More specifically, it uses *Levin's Kt complexity*, a modification of *Kolmogorov complexity*, to assign a mathematical value to the challenge put before an intelligence. (Kolmogorov complexity is a measure of the minimum computational resources required to define an object. However, it isn't computable, so an approximate value is derived using Levin's Kt complexity.)

The test is referred to as an "anytime test," because, as structured, it isn't dependent on time. A value can be derived from minimal interaction, with increasing accuracy as the time is increased and more challenges undertaken.

This approach allows us to tailor challenges to the level of an intelligence—be it animal, machine, or even, in theory, an alien—and assign a meaningful value to the result. An additional benefit of such an approach is that a subject isn't required to be able to understand language. This would allow us to test the intelligence of animals as well as those machines lacking natural language-processing capabilities.

Developing methods of interfacing with different types of intelligence will be a key problem. The intelligence tests might take the form of structured environments in which particular tasks are to be performed by the test subject. Using a series of observations, rewards, and actions, we could then calculate the aggregated complexity of these tasks to yield a useful value of performance.

The ideas that led to this approach have been developed by Hernández-Orallo and Dowe, as well as Shane Legg and Marcus Hutter, among others, and are based on the work of the pioneers of algorithmic information theory, Ray Solomonoff, Andrey Kolmogorov, Gregory Chaitin, and Chris Wallace.

The Future of Intelligence

Beyond assigning a single value of intelligence to nonhuman subjects, it should be possible to test and rate different factors of intelligence. Just as human beings differ in their balance of different intelligence factors, so, too, could machines. In fact, it can be argued that AIs could vary in their dis-

tribution of these far more than people do.

What if self-improving AI gives rise to a superintelligence? Should this occur, it seems more likely that many intelligences would ultimately develop, rather than a single monolithic one. Assuming these diverged in much the same way that biological organisms do through evolution, a broad variety could develop. Such tests could allow us to create a taxonomy of machine intelligences. Tests that gauge the "personality" of an AI based on its balance of intelligence factors could prove beneficial, if not life-saving.

An anytime test of universal intelligence offers a significant potential tool for futurists, scientists, and policy makers. By measuring machine intelligences accurately, it would improve our ability to track their development and make more-accurate trend analyses and projections. This would allow for a more-accurate assessment of where we've been and where we're going. Such knowledge would lead to better-informed scenarios, as would an improved understanding of existing and potential types of superintelligence.

As our world becomes increasingly filled with technological intelligence, it will serve us well to know exactly how smart our machines are and in what ways. Given that we try to measure almost every other aspect of our world, it seems only prudent that we accurately measure the intelligence of our machines, as well—especially since, by some projections, they're expected to surpass us in the coming decades.

This leads to the question: Is the day approaching when machines will debate whether or not human beings are truly intelligent?

Richard Yonck is a foresight analyst for Intelligent Future LLC. He speaks and consults about the future and emerging technologies and writes a futures blog at Intelligent-Future.com. He is also THE FUTURIST's contributing editor for Computing and AI. E-mail ryonck@intelligent-future.com. This article draws from his paper "Toward a Standard Metric of Machine Intelligence," published in *World Future Review's* special WorldFuture 2012 conference edition.

"What if self-improving AI gives rise to a superintelligence? Should this occur, it seems more likely that many intelligences would ultimately develop, rather than a single monolithic one."