Document downloaded from:

http://hdl.handle.net/10251/159605

This paper must be cited as:

Hernández-Orallo, J. (2019). Unbridled Mental Power. Nature Physics. 15(1). https://doi.org/10.1038/s41567-018-0388-1

The final publication is available at

https://doi.org/10.1038/s41567-018-0388-1

Copyright Nature Publishing Group

Additional Information

(*commissioned for the column "Measure for Measure", Nature Physics*, https://www.nature.com/nphys/articles?type=measure-for-measure)

748 WORDS

Intelligence without Measure

José Hernández-Orallo

There was a time horses were a major source of physical power. When the steam engine started to rival them, manufacturers wanted to know how many horses a particular engine would replace. James Watt soon realised how important these comparisons were, and conceived a new measure: the horsepower. From discussions with millwrights, a horsepower was estimated to be 33,000 ft-lbf min^{-1}. The measure was a great success. Yet, two centuries later, with horses rarely used for physical work, the horsepower has been replaced by the Watt.

Today, humans are still a major source of mental power, but *artificial intelligence* (AI) is starting to rival them. Again, one valuable comparison seems to be whether a particular AI system is more powerful than a *standard* human. And the recurrent question is when this will happen, under a malleable term: *human-level machine intelligence*. Readily, project management could lend us a measure for AI: the person-month.

Despite the stimulating analogies, there are many differences between physical and mental work. The early psychometricians pushed the analogy as far as they could, measuring *intelligence* as the capability of producing a particular kind of information processing work. However, psychometric measurement derives from human populations. In many of its forms, it just captures a deviation from the mean, but not an actual magnitude. No imperial foot for intelligence is there to be used as a ratio scale.

So back in the late 18th century, what Watt did was Copernican: he put horses in terms of universal physical measures —feet, pounds and minutes—, independently of any other horse. In Watt's time, the understanding of the physical world was sufficiently mature to realise that the power needed in a mill could be compared with the power needed to boil a pot of water.

In contrast, even today, there is nothing like a unit for mental power, independently of the human and independently of the task. In fact, the main problem for adapting psychometrics to AI is not the lack of a ratio scale, but its *populational* character. For obvious reasons, the notion of machine population in AI is thorny. Still, a bevy of AI competitions, benchmarks and platforms have been recently introduced[1]. Progress is measured in terms of performance on particular tasks, usually compared with some average human estimate. Cross-task comparison remains elusive, though, as many AI systems are specialised for a single task.

Some would say that *cognitive tasks* cannot be reduced to a limited number of capabilities, or even a single one. Different tasks could never be compared. But others would say that intelligence may have different manifestations and a complex structure, a phenomenon that is common in physics. From the performance of a system on a set of tasks, we could *predict* its performance on a different set of tasks.

The range between these two extremes—the importance of bias— is immanent within machine learning. From this background, Solomonoff's prediction theory[2] and Levin's universal heuristics[3] see Occam's razor as a bias that emerges from algorithmic information theory, a possible foundation for computational measures of intelligence. The elements are still insufficient, and are superficially very different from the dominant paradigm in AI today, deep learning. Still, they have more potential than any other current computational theory of intelligence. For instance, Levin's universal search makes it possible to define the difficulty of any inversion task —its required search *work*. From here, the capability of a system can be defined as an integral of performance over a range of difficulties. In this way, both difficulty and capability are measured on a ratio scale, with the same unit: *the logarithm of the number of computational steps*[4]. This unit is ultimately commensurate to *bits*, under the two terms of Levin's universal search.

This conceptually appealing formulation has some technical limitations. For instance, without the choice of a reference machine, Kolmogorov complexity and the logarithm of the number of computational steps will depend on constants. Interestingly, however, the answer to these limitations may lie in further linking computation and information to physics. Indeed, there must be bounds between mental power and physical energy, and discovering them may shed light on questions such as AI progress, intelligence growth, footprints on the environment and the effect of quantum computing on AI.

By seeking the units of mental power and linking them to physical units, we may look eccentric from the thriving perspective of an unbridled AI field. Like Watt two centuries ago, sometimes we have to put the cart before the horse.

JOSÉ HERNÁNDEZ-ORALLO is Professor in the Department of Computer Science, Universitat Politècnica de València, and Visiting Fellow at the Leverhulme Centre for the Future of Intelligence, University of Cambridge. email: jorallo@dsic.upv.es

References

1. Castelvecchi, D. "Tech giants open virtual worlds to bevy of AI programs." *Nature News* 540, no. 7633: 32 (2016).
2. Solomonoff, R. J. "A formal theory of inductive inference. Part I." *Information and control* 7, no. 1: 1-22 (1964).
3. Levin, L. A. "Universal Heuristics: How Do Humans Solve "Unsolvable" Problems?" In Dowe, D.L. (Ed.) *Algorithmic probability and friends. Bayesian prediction and artificial intelligence*, pp. 53-54. Springer, Berlin, Heidelberg, (2013).
4. Hernández-Orallo, J. *The measure of all minds: evaluating natural and artificial intelligence*. Cambridge University Press, (2017).