José Hernández Orallo - The measure of all minds - 2017

Tom Rochette <tom.rochette@coreteks.org>

August 30, 2025 — 861fb9d0

0.1 Context

0.2 Learned in this study

0.3 Things to explore

- Can we devise universal cognitive tests?
- Can we have a formal definition of intelligence solely based on computational principles?
- How do you evaluate a human-machine hybrid?
- What is the minimum amount of intelligence required for general intelligence?
- What is the maximum limit to intelligence, subject to physical limits?
- Can you use the theory in this book to explain how a group of individuals (collective) can be less productive than the sum of its individuals?
- How can you compare two ML/DL/AI models according to the theory presented in this book?

1 Overview

2 Notes

2.1 Chapter 1 - Extended Nature

2.1.1 1.1 Face the diversity

- Several groups
  - Computers
  - Cognitively enhanced organisms
  - Biologically enhanced computers
  - (Hybrid) collectives
  - Minimal or rare cognition

2.1.2 1.2 The machine kingdom

- No matter how wild and diverse life may be, it is constrained by the rules of evolution and natural selection
- The machine kingdom is the set of all interactive systems taking inputs and producing outputs, possibly asynchronously, through interfaces, bodies, sensors and actuators, etc.
- The physical Church-Turing thesis: every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means
- The same algorithm running in a faster computer is, in terms of the machine kingdom, a different machine
- We will generically refer to the elements of the machine kingdom as agents or subjects, especially when confronted with an environment, a task or a test

- Our fixation on the behaviour of all interactive systems derives from our interest in the measurement of what systems do
- Our goal is to measure and classify them (machines) in terms of their behavioral features. This is what we call universal psychometrics
- Universal psychometrics is the analysis and development of measurement tools for the evaluation of behavioural features in the machine kingdom, including cognitive abilities and personality traits
- Many hurdles had to be overcome, such as how to perform tests without the use of language, how to make animals focus on a task and how to choose the right interface and rewards
- In practice, artificial intelligence uses specialised tests for each particular task

2.1.3 1.3 The space of behavioural features

- A behavioural feature is an abstract property, characteristic or construct about the elements of the machine kingdom that can be inferred by observation and interaction
- To characterise the elements in the machine kingdom we will mostly focus on those behavioral features that are most stable: personality traits and cognitive abilities
- Personality traits, in particular, represents "consistent patterns of behavior, especially expressive or stylistic behavior"
- A personality trait is a behavioural feature of an interactive system in the machine kingdom that captures a facet of its character or temperament, describing what the system tends to do
- Cognitive abilities are properties that describe what subjects are able to do
- Cognitive abilities are usually gradient features, i.e., the more the better
- A cognitive task is an interactive series of stimuli that allows for different observable behaviours on the subject. It is cognitive as far as performance is involved, and its interface can be changed or virtualised without affecting the essence of the process
- A cognitive ability is a gradient property of an interactive system in the machine kingdom that allows the system to perform well in a class of cognitive tasks

2.1.4 1.4 Psychometric profiles and intelligence

- Adaptation is usually performed through specialisation
- If systems may perform well by innate (programmed) behaviour or by acquired (learnt) behaviour, should we evaluate whether subjects are able to do some tasks or whether they are able to learn to do some tasks?
- Does it make sense to evaluate whether a system can learn a single task instance or a wide range of tasks?
- What we are truly aiming at is a representative set of behavioural features, such that the elements in the machine kingdom can be characterised with them
- A psychometric profile, given a set of behavioural features, is the set of measured values for a particular subject

2.1.5 1.5 An opportune venture

- Black-box approach: A major aspiration of universal psychometrics is to design measurement tools that work when the only information we have available is the behaviour of a previously unknown system
- We should talk about superabilities and ask whether those can be measured beyond the human level

3 See also

4 References