A Universal Measure of Intelligence for Artificial Agents

Shane Legg
IDSIA — Switzerland
shane@idsia.ch

Marcus Hutter
IDSIA — Switzerland
marcus@idsia.ch

1 The concept of intelligence

A fundamental difficulty in artificial intelligence is that nobody really knows what intelligence is, especially for artificial systems which may have senses, environments, motivations and cognitive capacities which are very different to our own.

If we look to definitions of human intelligence given by experts, we see that although there is no consensus, most views cluster around a few common perspectives and share many key features:

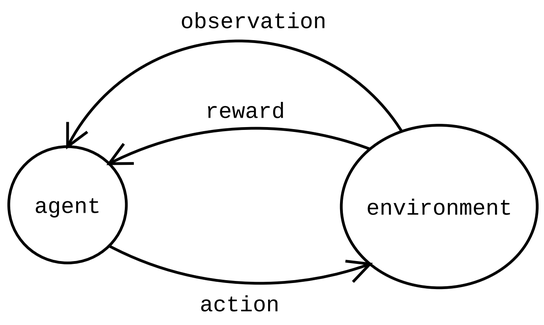
- intelligence is a property of an *agent*
- the agent interacts with an external *environment*
- related to success with respect to some *goal*
- the environment is not fully known to the agent

The last condition implies that the agent must be able to learn and adapt to unknown environments based on experience. This gives us our informal definition of intelligence:

> *Intelligence measures an agent's general ability to achieve goals in a wide range of environments.*

Here we will try to formalise this view of intelligence.

2 A formal framework



We use reinforcement learning as our formal framework as it is both simple and extremely general. We call the signals sent from the agent to the environment *actions*, and the signals sent back *perceptions*. The perceptions are divided into two parts: A signal that indicates the agent's success, called the *reward*, and a non-reward part called the *observation*.

The observation, reward and action symbols being sent between the agent and the environment are denoted by lower case variables o, r and a. They are indexed in the order in which they occur, thus a_3 is the agent's third action. This process of interaction produces an increasing history of observations, rewards and actions, $o_1 r_1 a_1 o_2 r_2 a_2 o_3 r_3 a_3 o_4 \ldots$.

The agent is a function, π, which takes the current history as input and chooses the next action as output. We represent this as a probability measure over actions conditioned on the current history, $\pi(a_3|o_1 r_1 a_1 o_2 r_2)$. The internal workings of the agent are left unspecified. The environment, μ, is similarly defined: $\mu(o_k r_k|o_1 r_1 a_1 o_2 r_2 a_2 \ldots o_{k-1} r_{k-1} a_{k-1})$.

As the reward is generated by the environment, the agent's goal is implicitly defined by the environment. Thus to test an agent in any given way it is sufficient to define its environment.

The agent must try to maximise the total reward it receives over time. The standard way of expressing this is to weight the future reward at time i by a factor γ_i,

$$V^{\pi\mu} := \mathbf{E}\left(\sum_{i=1}^{\infty} \gamma_i r_i\right),$$

where r_i is the reward in cycle i of a given history, and the expected value is taken over all possible interaction histories of π and μ. The choice of γ_i is a subtle issue that controls how greedy or far sighted the agent should be. Here we use the near-harmonic $\gamma_i := 1/i^2$ as this produces an agent with increasing farsightedness of the order of its current age [Hutter2004].

As we desire an extremely general definition of intelligence for arbitrary systems, our space of environments should be as large as possible. An obvious choice is the space of all probability measures, however this causes serious problems as we cannot even describe some of these measures in a finite way.

The solution is to require that the measures which represent the environments are computable. This allows for an infinite space of possible environments with no bound on their complexity. It also permits environments which are non-deterministic as it is only their distributions which need to be computable. This space, denoted E, appears to be the largest useful space of environments.

3 A formal measure of agent intelligence

We want to compute the general performance of an agent in unknown environments. As there are an infinite number of environments in our set E, we cannot simply take a uniform distribution over them.

If we consider the agent's perspective on the problem, this is the same as asking: Given several different hypotheses which are consistent with the data, which hypothesis should be considered the most likely? This is a standard problem in inductive inference for which the usual solution is to invoke Occam's razor:

> *Given multiple hypotheses which are consistent with the data, the simplest should be preferred.*

As this is generally considered the most intelligent thing to do, we should test agents in such a way that they are, at least on average, rewarded for correctly applying Occam's razor. That is, test in such a way that simpler environments really are more likely. In our framework this means that our a priori distribution over environments should be weighted towards simpler environments. However to do this we need a way to measure the complexity of environments.

As each environment is described by a computable measure, one way of measuring the complexity of an environment is by taking its Kolmogorov complexity. If $\mathcal{U}$ is a prefix-free universal Turing machine then the Kolmogorov complexity of an environment μ is the length of the shortest program on $\mathcal{U}$ that computes μ,

$$K(\mu) := \min_p \{l(p) : \mathcal{U}(p) = \mu\}.$$

Unfortunately, K is not computable and is provably difficult to approximate. For the purposes of Occam's razor, it also seems philosophically unnatural to consider short programs which require an enormous amount of time to compute to be "simple". We can address both of these problems by using a notion of complexity that takes execution time into account, such as Kt complexity [Levin1973],

$$Kt(\mu) := \min_p \{l(p) + \log t(p) : \mathcal{U}(p) = \mu\}$$

where $t(p)$ is the number of steps required to compute μ on $\mathcal{U}$. This gives us a computable distribution $2^{-Kt(\mu)}$ over our space of possible environments which is consistent with the notion that very simple algorithms should be short and fast to compute. Another alternative based on similar ideas is the Speed Prior [Schmidhuber2002].

We can now define the *universal intelligence* of an agent π to simply be its expected performance when faced with an unknown environment sampled from this distribution,

$$\Upsilon(\pi) := \sum_{\mu \in E} 2^{-Kt(\mu)} V^{\pi\mu}.$$

4 Properties of universal agent intelligence

This universal measure of intelligence for artificial agents has many important properties:

Formalises common informal definitions It is clear by construction that universal intelligence measures the general ability of an agent to perform well in a very wide range of environments, similar to many informal definitions.

Very general The definition places no restrictions on the internal workings of the agent; it only requires that the agent is capable of generating output and receiving input which includes a reward signal.

Non-anthropocentric Universal intelligence is based on fundamentals of information and computation theory. In contrast, other tests such as the Turing test are largely a measure of a machine's "humanness", rather than its intelligence.

Incorporates Occam's razor In this respect it is similar to intelligence tests for humans which usually define the "correct" answer to a question to be the simplest consistent with the given information.

Spans low to super intelligence Universal intelligence spans simple adaptive agents right up to super intelligent agents, unlike the pass-fail Turing test which is useful only for agents with near human intelligence.

Practically meaningful A high value of universal intelligence would imply that an agent was able to perform well in many environments. Such a machine would obviously be of large practical significance.

By considering $V^{\pi\mu}$ for a number of basic environments, such as small MDPs, and agents with simple but very general optimisation strategies, it is clear that Υ correctly orders the relative intelligence of these agents in a natural way. If we consider a highly specialised agent, for example IBM's DeepBlue chess super computer, then we can see that this agent will be ineffective outside of one very specific environment, and thus would have a very low universal intelligence value. This is consistent with our view of intelligence as being a highly adaptable and general ability.

The definition given here can be seen to be a simplified version of the Intelligence Order Relation (IOR) [Hutter2004]. By definition, the maximal agent with respect to this order relation is AIXI, and with minor technical adjustments, AIXI would also be maximal with respect to Υ. AIXI has been shown to have many optimality properties, including Pareto optimality and the ability to be self-optimising in environments in which this is at all possible. This demonstrates the power of agents with very high universal intelligence.

The only related work to ours is the C-Test, see for example [Hernández-Orallo2000]. While we have defined a fully interactive test, the C-Test is a static sequence prediction test which always ensures that each question has an unambiguous answer, like a standard IQ test. We believe that these are unrealistic and unnecessary assumptions. The C-Test was able to compute a number of usable test problems which were shown to correlate with real IQ test scores when used on humans.

References

[Hernández-Orallo2000] J. Hernández-Orallo. Beyond the Turing test. *Journal of Logic, Language and Information*, 9(4):447–466, 2000.

[Hutter2004] M. Hutter. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Springer, Berlin, 2004.

[Levin1973] L. A. Levin. Universal sequential search problems. *Problems of Information Trans*, 9:265–266, 1973.

[Schmidhuber2002] J. Schmidhuber. The Speed Prior: A new simplicity measure yielding near-optimal computable predictions. In *Proc. COLT 2002*, Lecture Notes in Artificial Intelligence, pages 216–228, July 2002. Springer.