Towards a Theory of General Intelligence

Abel Torres Montoya, DATAVERAS

Problem Overview

Despite recent progress in ML we still lack even a rudimentary theory of intelligence for general and efficient problem solving.

This work introduces a framework for implementing intelligence based on a small set of principles. This method have common points with several existing proposals: it is based on a *master algorithm*, involving an *energy equilibrium*, there are *logic-based* mechanisms expressed in *network dynamics*, the system is *self-motivated* and *open-ended,* and, finally, *feedforward* reactive mechanisms coexist with *generative abstracted representations*.

'General' Intelligence

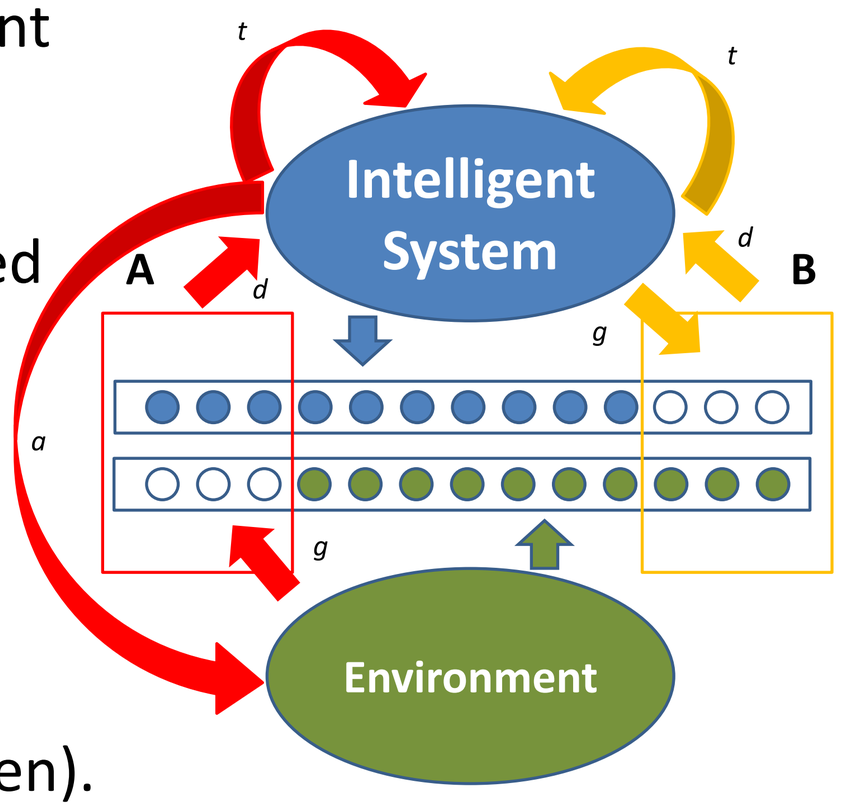
Ambiguity and stereotypes in the definition of intelligence are hampering the creation of a proper frame for developing an AGI theory.

As we shown in the proposed framework:

- There is no 'general' vs 'narrow' dilemma: intelligence is intrinsically general and every implementation is intrinsically constrained (i.e. narrowed)
- Value assignment for decision-making is part of the implementation-based 'narrow' functionality
- Intelligence is indivisible from following properties: unsupervised/lifelong/zero-shot learning, abstraction/imagination, counterfactual reasoning

A New Paradigm of Intelligence

At high level intelligence is a system to mimic the environment used by the system/organism to learn (unknown input) or to achieve a goal (unmatched internal representation). That is achieved by keeping a balance with the input using abstracted generative representations. Abstractions are created using innate priors (e.g. the Hebbian rule captures the spatio-temporal consistencies in the input).



In the figure an open-ended, equilibrium-based mechanism unifies learning (yellow) and problem solving (red). Internal representations (blue) should match input from context (green).

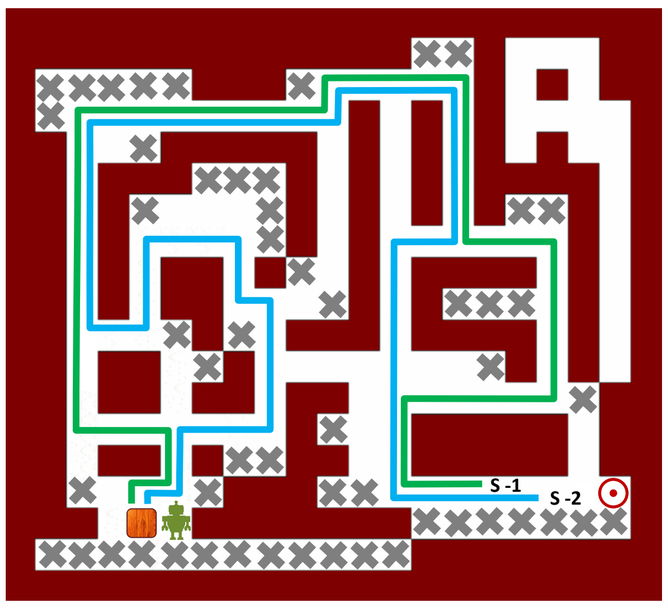
Logic & Reasoning in a Network

This paradigm requires removing the already known parts of the input. This is achieved using inhibition operations. It is shown that simple inhibition mechanisms allow creating open-ended reasoning and counterfactuals.

Thanks to logic capabilities it is possible to do solution space coverage, detecting when a solution is impossible or going over the space of possible solutions. Parallels between the described dynamics and neuroscience evidence are discussed.

From Theory to Prototyping

Current efforts are focused on prototyping *ad-hoc reasoning* in interactive, large-scale exploratory scenarios (e.g. user modifiable Sokoban challenges).



Extension to robotics and protein design is envisioned.