

AGI v0.2.0

Tom Rochette <tom.rochette@coreteks.org>

January 10, 2017 — [83cc42a4](#)

0.1 Learned in this study

0.2 Things to explore

- Attention/focus-based systems ala EURISKO
 - Simulation/Experiments
- Learning seems to be trial and error at first (randomly) until a couple of “provably” working techniques are developed, which are then iterated on
- Thought experiments
 - How would I go about learning to do X?
 - How would a group of individuals go about learning to do X?

1 Overview

There are many different routes one might want to look into.

1.1 Mathematical and formal approach

We look into mathematical models and formal approaches such as AIXI/OOPS/Gödel machines where what we are doing is searching for programs using a proof searcher and a program runner. This approach thus requires the implementation of a proof searcher as well as some system that can execute the programs that have been found. The AGI will also need some form of reward stream or evaluation function which will allow it to determine if a newer program is superior to the current running program. This reward stream has the purpose of guiding the AGI toward a given goal and as such, the AGI will optimize itself to get to that goal as fast as possible. The difficulty of this approach is figuring out how to reward the AGI properly so that it does significant work or in other words, determine an appropriate reward evaluation function.

1.1.1 What is missing?

- The seed proof searcher
- The evaluation function

1.2 Tasks system approach

This approach is to build some sort of task runner system, where perception is one of many tasks the system has to deal with. Other tasks are internal decisions, evaluations, prioritization, etc. In this approach, we have to implement most of the internals of the AGI. The nice thing about this approach is that we can program it in such a way that we may inspect and understand it, however it also means that it will most likely be limited to how it was designed and will make it difficult to evolve on its own.

1.2.1 What is missing?

- The necessary functions

- The code for all these functions
- How to hook everything together

1.3 Neural network approach

An approach similar to the task system approach could be to use many specialized neural networks together in order to accomplish specific tasks. Each neural network would be trained for a single purpose and a general neural network would be tasked with determining which specialized neural network should process the task at hand.

Using an approach similar to the one presented by AlphaGo, this AGI would use a policy network and a value network in order to drive its decisions. The policy network would be used to evaluate the potential actions based on previous experiences (and their probabilities) while the value network would help the policy network decide toward which direction it should go.

1.3.1 What is missing?

- How to feed the network
- How to train the network
- What is the architecture of the network(s)?

1.4 (Software) Brain replication approach

This approach is very much related to the neural network approach. The idea is to reverse engineer how the human body and the brain work in order to accomplish goals.

It makes sense to start first with how human embryos grow into humans simply through the use of the genetic code that was provided by its parents. This code is responsible for the formation of all the parts of the body through the consumption of energy and the obtention of the appropriate “construction” components from its host, its mother. This code could effectively be considered to be the bootstrapping code of the AGI.

But the information provided by the “growth” code may not be useful in itself to the purpose of the AGI we are searching for: one that can accomplish tasks given a goal. It may only be of interest to understand how the fully grown neural system works in concert with the different modalities of the body.

The interesting aspect of this approach is that it uses us, as human, as a proof that it is possible to produce such machine. Given the proper nutritive medium and the embryo, a host can birth a child that will be able to accomplish goals and have a life of its own.

1.4.1 What is missing?

1.5 (Hardware) Brain replication approach

Similarly to the software brain replication approach, the hardware brain approach is based on the idea of reproducing the same processing architecture of the nervous system of the human body.

Here, there are many areas of research, but many could be considered ethically discutable.

- Transfer of deceased brain in a vat: Keep the brain alive in some sort of vat while it is fed nutrient and oxygen like the blood would. Every sensory inputs and motor outputs would be linked to some electronic system which would then convert these signals into either a simulation of the real world or virtual world, or to a remotely-controlled robot.
- Brain culture: Through a better understanding of embryonics and genetics, it becomes possible to “grow” brains in a culture, similar to how we can grow cells (or we are currently attempting to grow organs). In this case there are many challenges: how to grow a brain? what are the requirements? how will we interface with it? will it act exactly like a “regular” brain? is an out of body brain going to develop the same way?

An interesting aspect of this idea might be that if we truly have specialized regions of the brain, we may be able to “breed” specialized brains for vision, audition and processing of sensory inputs.

- Live brain: It may be possible to interface with existing brain and use them without interfering with their current work. This would allow us to execute computation on existing “surfaces”: people.

In this particular approach, the issues are numerous (some of them ethically discutable):

- We need a better understanding of how the brain works
 - Composition, interface, assembly, structure, organization, etc.
- We need a better understanding how of cells become specific part of the brain/specialize
- We need to learn to “cultivate” organs as complex as the brain
- We need to attempt to preserve the brain of deceased individual

2 See also

3 References