AGI v0.1.0

Tom Rochette <tom.rochette@coreteks.org>

June 23, 2018 — 0fa579d5

0.1 Learned in this study

0.2 Things to explore

- How can we discover that a vector of data makes more sense if structured as a MxN matrix?

1 Overview

The brain can be seen as a single function machine. It receives inputs, processes them, and finally return some instructions as to what to do next. However, what this function does is a complete mystery to all of us.

The input -> process -> output model is very easily understood, however each part in itself is filled with questions: how is the input formatted? how much data do the different channels contain? is all data processed in the same way? is data processed the same way every time? are there any channels with more priority than others? how is the same information processed faster the next time it is observed? how does the brain know how to make use of its available outputs to act on its environment? how does it know what outputs are available? what processes are used to make recalling faster/more accurate when needed? what are the most barebone requirements for each of the components?

The following document will attempt to answer those questions by providing an architecture/framework which attempts to put together all of the existing knowledge about AGI. Obviously, since AGI development is a new field, there are many schools of thought as to how one should be built, and how I suggest it be built is one of many ways we *might* end up building one. Don't take any of what I say for granted but as a source of ideas for your own architecture.

Feel free to drop me a comment or any questions down below (at the end of the article).

2 Layers

```
graph TD;
    0[Start]
    0 --> 1[Main loop]
    1 --> 2[Input]
    2 --> 3[Process]
    3 --> 4[Output]
    4 --> 5[End loop]
    5 --> 1
    5 --> 6[Done]
```

3 Pre-requisites

3.1 Layer 1 - Storage

Storage is the primary source of knowledge and experience which can later on be used for action-reaction type of response to stimuli. Without storage, we can at best give an initial set of immutable action-reaction instructions to the AGI (similar to the code of a program which cannot be modified).

4 Input

4.1 Layer 2 - Processing/Pattern detection

4.1.1 Layer 2.1 - Data streams

The main source of information we receive is based on time. We see one thing after the other and thus it is important to be able to process what we perceive in the same fashion as computer process data streams (as opposed to batches of data).

Data stream processing will interact a lot with the storage layer as it where the observed and extracted patterns will be stored. Furthermore, much like a computer the AGI will need to have various levels of caches available to it in order to store frequently/recently used patterns for fast access/retrieval.

What appears to make the most sense in term of pattern detection is the use of a hierarchical pattern detection algorithm, where what is being “detected” goes through some sort of binary tree. Thus, pattern detection would be executed similarly to binary search (in its simplest form). If it is however possible to compare the stimulus against multiple patterns at once, we may compare the process to something like a B-tree search. The means through which the tree is ordered may depend on various factors, such as how often this pattern has been observed, or how recently it has been observed (think of a splay tree adjusting itself based on search).

4.1.2 Layer 2.2 - Data batches

Send data to the data stream layer

For the purpose of being able to process data that is given to us by other AGI agent, it will be useful to be able to go through this information in a batch. Information can however be transferred and processed as a data stream for which the timestamp is already provided.

4.2 Layer X - Input filtering/Attention

(might be mergeable into the processing/pattern detection layer)

As computing power increases, it may be possible to reduce the amount of data filtered in favor of processing more data. However, as more data gets processed, it is still important to be able to extract the meaningful information out of the noise.

5 Process

5.1 Layer 3 - Decision

Once we have recorded and processed information from the environment, the next step is to decide if we’ll want to do anything with it.

5.1.1 Layer 3.1 - Simple reflex/Automated responses

5.1.2 Layer 3.2 - Model-based/Analysis/Prediction

5.1.3 Layer 3.3 - Goal-based

5.1.4 Layer 3.4 - Utility-based

5.2 Layer 4 - Plan execution

this will need to be divided into many more components or the plan execution should be short lived and work off feedback (running through the whole process again and again, replanning as required)

Given a decision by the previous layer, the plan execution layer is tasked with sending the appropriate signal to the required systems (may it be to store some information for later recall, activate motors to move a limb at the appropriate location, etc.).

6 Specialized layers

6.1 Layer N - Natural language processing

6.2 Layer N.1 - Understanding

6.3 Layer N.2 - Meta data recording

Remember where the information came from (conversation, document, who were involved)

6.4 Layer M - Logic

6.5 Layer A - Inference

For the AGI to be able to do about anything, it will have to be able to make inferences such that if event X happens, then event Y happens shortly after as well, all the time, then these two events are most likely correlated. It is also critical for the AGI to be able to do inference if it wants to be able to recognize that its actions are having some result on the environment it is in.

7 Arrival of new data

1. The data is taken as a single unit. An **Identifier** is produced to uniquely identify it.
2. An **OrderedSet** is produced. It is uniquely identified by the **Identifier** that was just produced.
3. The data received is stored in a single **Chunk**. That **Chunk** is pointed to by the **OrderedSet** that was produced in step 2.

8 Arrival of new data with existing and overlapping data

9 Output

10 Question answering

10.1 How is the input formatted?

Let's look at a real world example: A developer is interacting with a database visualization software in order to understand the schema of a database.

Conceptually, the database is composed of multiple tables, which themselves are composed of columns and their own specification (type, length, nullable, etc.). However this information is not accessed directly by the developer. Instead, it has to go through his vision system (retina, optic nerve, optic chiasma, optic tract, lateral geniculate body, optic radiation, visual cortex, visual association cortex) and then processed into high level concepts such as “database”, “table”, “column”, “type” and so on.

What I want to point out of this example is that the sensory modalities are defined once and for all. After the signal has been received (by the retina in our example), this raw signal has to be processed and converted into something our conscious mind is able to deal with.

If we make the parallel with computers, then it means that our most basic sensory modality format is most likely going to be the stream/packet of bits. Then it is up to the systems related to this sensory modality to make sense of the data that was received.

The question at this point then becomes, what are the sensory modalities that make sense? If we built a robot for instance, then the webcam could be considered as the visual and auditory systems. Taste and smell could be implemented through some form of chemical analyzer (if necessary). Finally, motor actuators would serve as output to allow the robot to move within its environment.

However the robot may also be virtual. In this case, a stream of data could be sent to it and “tagged” as visual, or auditory, and then the purpose of the algorithm would be to make sense of it (e.g. discover the data is 2D (an MxN matrix), its size/resolution, how color are represented (if there are any), etc.).

In short, what this means is that as long as the input contains enough information for the AGI to process, then it is part of the AGI task to figure out what it contains. However, once the format has been established, then newer information can now use this structure to build upon.

Another question that comes out of this discussion is “what is the impact of having a single input channel instead of many?” As the visual and auditory systems are separate in humans, we are to wonder what is the purpose of the separate channels. Is it simply because some information simply cannot be processed through the given modalities? One cannot hear through his eyes nor see through its ears. Thus the different modalities we’ve acquired are potentially adaptation to different form of data that we could not process with our sensory modalities at the time. A simple explanation as to why it is this way is that having multiple “streams” multiplexed into one adds additional burden to the processing of information, namely demultiplexing the inputs. It is far simpler and cleaner architecturally speaking to have highly cohesive data going into a function than it is having the function deal with an unrelated structure to extract the information it wants to process.

Another interesting bit of information is that all signals within the brain are transmitted through neurons and thus have to be converted into a chemical/electrical signal at some point. What this means is that whatever the input format is, it will have to be converted in a 0/1 equivalent.

10.2 How much data do the different channels contain?

10.3 Is all data processed in the same way?

10.4 Is data processed the same way every time?

10.5 Are there any channels with more priority than others?

10.6 How is the same information processed faster the next time it is observed?

10.7 How does the brain know how to make use of its available outputs to act on its environment?

10.8 How does it know what outputs are available?

10.9 What processes are used to make recalling faster/more accurate when needed?

10.10 What are the most barebone requirements for each of the components?