An open source Artificial Intelligence research and development project.

While many artificial intelligence projects work on creating a neural network to simulate a brain, I would like to propose looking into a psychological perspective. While building a simulated structure of the brain is important, it is only really so for the neuroplasticity and organization of the brain. It is the platform that the mind sits on, but a brain without the mind is no better than a computer without an operating system.

Understanding the psychology of the human mind would be more beneficial to AI development, because this would tap more into understanding the mind of a human, not the brain. By simulating how the mind works, you are building that operating system, giving the program human like functions. The means of doing that is creating the most basal elements of psychology such as the understanding of the process of thought and action, the meanings of words, and being able to understand the world through comparison and contrast.

The Process of Thought

The process of thought and action can be broken down into these words: Stimulus, Observation, Thought, Plan, Action, Verification. I will go over these steps in detail.

Stimulus

This is the external or internal signals that get picked up by the individual. Essentially, they are inputs. These inputs however are also linked to the output at the end of the cycle of verification, since a result can become the stimulus, causing another chain of thought. An example that this phenomenon is most commonly seen is in research, where while finalizing their initial experiment, more questions gets generated from the original question.

Observation

This step is important, because this constitutes as the attention that the AI gives to the stimulus, analyzing and using methods to compare and contrast to memories (past data) and other stimuli. The methods of comparison are through statistical analysis and other forms of comparison studies. This is where attributes gets attached to objects, and differences through pattern recognition are made. This stage is done for every stimulus, and should happen quickly and automatically. Some observations require longer than others, but this gets determined if it is called for in the Plan stage, not before that. All data then gets passed for approval to the thought stage.

Thought

This step is a priority organizer, that looks at the information provided and decides to either disregard it (do nothing), go straight to action (fast reaction or low priority), or to the planning stage (high priority or requires more information). Most thoughts that occur in the human mind are disregarded, as there are many stimuli that are unimportant. Threshold levels need to be set to disregard noise or useless data.

Plan

Planning is the gathering of information, and running of simulations to get the best result for a course of action. This step will involve looping back through the Stimulus, Observation, and Thought steps, while also simulating possible outcomes if the individual takes action. Humans call this thought experiments. The result of this step should be an ordered list of instructions that the AI will take, which will get passed to the Action state in this fashion.

Action

This is the actual output that the AI will have, or the actions that it will take. This should be the end result of the Stimulus, Observation, Thought and Plan steps. No further calculation is done in this step, for it should already be finished.

Verification

This step is important, for after the action is performed, the AI verifies the response and compares the actual result to the expected result. If satisfied, this gets recorded to memory, allowing the AI to access it easier in the future. This is how it will learn new things, and gain new experiences.

The Human Eye - Functionality and Translation to an AI

The way a human is able to process an image is twofold. The first has to do with the mental capabilities of the occipital lobe, and the other has to do largely how the eye is set up, and what images are actually passed to the brain for processing. The amount of data is very minimal, since our vision is extremely different than that of a webcam.

The first obvious difference is that humans have 2 eyes, creating stereo vision which helps with depth perception. While you can just add two webcams, what must be discussed is the Depth of field of the eye (caused by the iris), and your actual clear vision compared to peripheral vision.

Depth of field is a very interesting "censor" for the brain, as whatever distance is looked at is focused on, while the rest becomes blurry. The easiest method to create this Depth of field is to have a telescoping lens that adjusts the focus mechanically, which is essentially how the eyes work. This can also be done post-image processing with libraries such as OpenCV and others, but will most likely require a sensor that can tell depth such as infrared or ultrasonic. The other method is using trigonometry, as there are programs (especially mobile apps) which calculate it this way. Whatever method, depth perception is needed to censor out background and foreground noise from the processing.

It is known that the human eye can only focus on one thing at a time, but the amount of view that the eye can see is surprising. 5% is perfectly clear , which is the very center of your focus. Approximately 18% of the vision surrounding the center is considered your clear vision. This is actually the total area that your brain processes, and this is where your focus lies. Everything else is peripheral vision, getting progressively blurrier the farther away from the center you go.

So in total, the amount of visual information that gets passed and processed critically to your brain is actually just 18% of the center image at a specific Depth of Field. Everything else, your peripheral vision, is glanced at by your brain to check for anomalies or significant objects. These consist of contrasting colors, fast moving detections, suddenly close objects, or anything that is considered "new" and "odd". When this is found, we turn our eyes, bringing it into that 18% Depth of Field image to actually process the information.

Other things worth mentioning is the how Field of View works with human vision. Quoting from Wikipedia (not a credible source, though it checks out based on personal observations) "The approximate field of view of an individual human eye is 95° away from the nose, 75° downward, 60° toward the nose, and 60° upward, allowing humans to have an almost 180-degree forward-facing horizontal field of view." This is made possible from the morphology of the face.

To break down the quote into steps, first it mentions "95° away from the nose ", which is because there is no obstruction on the left and right of your eyes. "75° downward " is this degree since your cheekbones blocks the rest of the vision. "60° toward the nose " is caused since your nose and nose bridge obstructs the rest of the vision, and the same with your eyebrow ridge for "60° upward ". This has its purpose also, as if we were able to see full 95% for all sides, there would be a lot of data flowing into our brain, as well as conflicting angles due to the fact that we have a left and right eye working together. If our left eye saw everything our right eye can see, and vice versa, we would most likely become confused with the conflicting data coming into our brain.

Words and the Mind

While the process of thought is essential, words are equally as important to the mind. It is known that human beings think with words, and the mind gets developed through understanding the words that we communicate with. We come to know these words through the characteristics attached to them, and these attributes are found from the process of thought mentioned above. This is how words resonate with us, because these characteristics connect with the senses, and invoke stimuli as if that word was manifested as a real entity.

To understand how this happens, we need to establish grammar and break down how sentences are constructed, and the purpose of each word within a sentence. If we do this, we can then pass a sentence to the AI, which will break it down and analyze what was said, converting the information presented as stimuli, and through the process of thought the information will become incorporated.

As it turns out, sentences are constructed in a very logical manner, and the direct definitions of how grammer is used can be translated almost directly into code for the AI. For example, lets take the sentence, "The cat is grey". Breaking down each word, the word "THE" is a declaritive word, almost like preparing for an object or directive in the next word. Following "THE" is "CAT". In the beginning, the AI (just like a baby) will not know what "CAT" is, therefore it will be left as an open definition, or perhaps a class element. The word "IS" states that "CAT" contains a property, which will be set in the next word. Finally, the word "GREY" is another word that is unknown to the AI (until learned), but it knows to place "GREY" as a property of "CAT".

The AI now knows that "CAT" is "GREY", so if we feed a statement such as, "This is a cat." (and show a picture of it), The AI will take this image and associate it to the class CAT. As you further talk about cats, and show images and behaviors of cats and so forth, the AI will keep adding these properties to the class CAT until it has a rather complex idea of what a cat is, and this is essentially how humans learn. We learn through observation and comparison.

But still, in the previous statement, we know what a cat is, but not what grey is. If we now pass the statement "Grey is a color", the AI will make the class GREY inherit the class COLOR. Through this inheritance system, we can organize the AI's thoughs so that it knows what belongs where, and also help with memory reduction since it can just inherit the properties of the large group. (This inheritance system also applies to the cat example above.)

This documentation is created by George Milord∴