COMMON SENSE REASONING

Than Lwin Aung

Senses

- Our consciousness is actually a black box; it is an isolated nervous system which has no actual connections with real world.

- The only thing which our consciousness does is making sense from all the information of our sensations from our sensory organs.

- How our consciousness can make sense from such diverse information from different senses?

So what makes sense?

Common Sense

- Common Sense, although pretty much common to humans, is the one of the most elusive concept in Humanity.

- In special areas, there are Artificial Intelligence Systems, which is comparable or even exceed human's capability.

- However, when it comes to "Common Sense", which seems common enough to humans, it is actually not common at all.

So what really is common sense?

Keywords

- Senses, Experiences, Knowledge, Reason, Emotion, Instincts, Habits, Wisdom

Purpose

- Philosophically speaking, many people have speculated on "the reason of our existence" or rather "the meanings of life".

- Some school of thoughts, such as Existentialism, believe that our life (existence) has no substantial meanings, while others believe our life serves some purpose.

- However, from the Scientific Perspective, our sole purpose of existence seems to be:

> How to **continue** to **exist** in an **unknown** environment with **limited** resources?

Sensibility

- From this interpretation, what makes senses is always in align with our purpose of existence i.e., the continuous existence in the unknown environment with the limited available resources.

- In other words, our existence is always bounded by two fundamental factors: **the unknown environment** and **the limited resources**.

- And thus, we always have two conflicting constraints: exploring the unknown as much as we can (**Exploration**) and saving the resources as much as we can (**Exploitation**).

- What is sensible is:

Exploitation   Sensible   Exploration

$\longleftrightarrow$
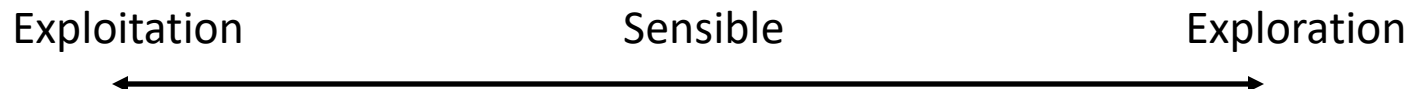
Strategy

- Formally, a Strategy is **allocation** of **right resources** to the **right place** at the **right time**.

- In other words, a Strategy is concerned with Exploitation Problem.

- However, how could we know if a strategy is good or bad?

- That's why Strategy alone is not enough to solve Exploitation Problem.

Decision

- Formally, a Decision is selecting **the Best Strategy** from a Set of Available Strategies given particular Conditions and Constraints.

- Unlike Strategy, a Decision involves some Exploration.

- To determine, what strategy is the best, we have to explore what strategy gives the best results. Otherwise, a decision cannot be reached.

- However, a Decision alone cannot solve the exploration problems.

- A Decision can be formalized with Decision Theory; for example, Game Theory is one of the Decision Theories, which assumes Decision Making is a Game.

Plan

- Formally, a Plan is a **Sequence of Decisions** from the **Initial State** to the **Final State**.

- In other words, a Plan can be assumed as a Sequential Decision Problem with Finite Steps (Decisions).

- Actually, we can see that Programming is actually a Planning Process.

- Any program (plan) will start from the Initial State and will eventually end in the Final State after Finite Steps (Decisions).

- Computability is concerned with "Planablity".

- What is not computable is what cannot transit from the Initial State to Final State within Finite Steps.

Policy

- Formally, a Policy is an **adaptable Plan**, without changing the Goal.

- Generally Speaking, a Plan is fixed Sequence of Decisions, which cannot be changed.

- A program (plan) cannot rewrite the steps itself after being complied.

- With a policy, an adaptable Plan, Programmability has become Learnablity (Adaptability) in which a program (plan) can change itself with respect to the dynamic environment.

- Therefore, a policy has become a learnable (adaptable) planning process, which can also solve some exploration problems.

- However, a policy still cannot change the Initial State and the Final (Goal) State; in other words, a policy does not involve "Goal Setting".

Reason

- Formally, a Reason involves exploring the Unknown based on **the Organization and Abstraction of Sensory Data**.

- Generally, sensory data is organized with temporal, spatial and similar relationship, thereby creating the "Knowledge Graph", which is the **abstraction of reality**.

- However, the "Knowledge Graph" itself is incomplete. That's why "Reason" is always incomplete.

- The incompleteness of "Reason" actually comes from the **absence** of "Risk".

- As we stated earlier, the primary purpose to explore the unknown itself is to increase our chance of survival: nothing more or nothing less.

Risk

- When exploring the Unknown, there is always **a chance of failure**.

- However much we are familiar with the conditions and constraints, a sudden change in some variables can cause major dynamic problems.

- Yet, "Pure Reason" does not have such assessment.

- For example, is $1 + 1 = 2$ is more risky than $y = 3x + 4$?

- It is also why "Reason" alone is never enough for assessing the "Risk".

- At any rate, the "Reason" must be aligned together with the "Risk".

Instinct

- Formally, Instinct is an **Reactive (Automatic) Risk Assessment**.

- There are two primary instincts: Survival Instincts, which assess the risk involving our survival, and Social Instincts, which assess the risk involving our social group.

- By "Reason" and "Instinct", it is possible to reduce the Risk while exploring the Unknown.

- "Reason" will enable us to recognize a "Person" in a picture; however, "Instinct" will assess how "Risky" that "Person" is by conveying some "Emotions".

- For now, AI has yet to incorporate "Instinct", but I am working on it right now.

Sentiment

- Formally, a Sentiment is a **Measurement of Risk**.

- It may sound a bit strange if we put it that way. However, our emotions are always centered around our risk of survival. We feel more "comfortable" when we are in "low risk" environment. As a rule of thumbs, no one will be smiling or laughing in risky (serious) environments. I think happiness is the measurement of risk.

- Therefore, I think sentiments play an important role to guide the AI as an intrinsic risk assessment.

- The sentiment triggers the instincts which determines our assessment of risk.

- Of course, the measure of risk as a sentiment is a complicated process, and I am working on it.

- However, is it possible to compute "Sentiments"?

Motive

- Formally, Motive is actually the **Process of Goal Setting**.

- Actually, what we call "**Desire**" is nothing more than setting goal.

- However, the ability to set Goal at will, given conditions and constraints, seems to be the **highest form of Intelligence**.

- In a sense, "Motive" controls "Instinct", "Sentiment", "Reason", and "Policy".

- How are a group of organic molecules able to set a Goal is still a mystery to me.

- At any rate, Artificial General Intelligence (AGI) should be complete when AI Agent will be able to set Goal at will.

Intelligence

Motive

Instincts	Reasons

Sentiments	Recognitions

Data
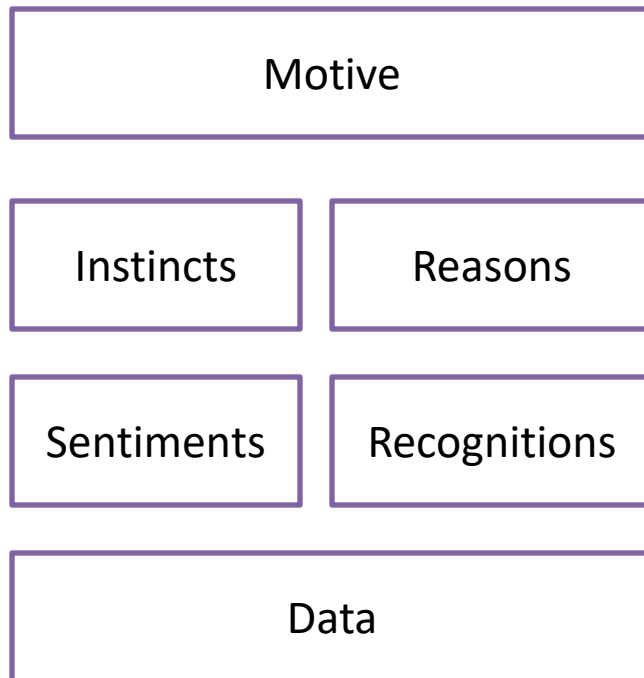
Reasons are primarily concerned with Knowledge Graph (Knowledge Model) as to Adaptable Planning and Decision, while Instincts are primarily concerned with Risk (Risk Model) which is associated with Knowledge.

Motive will set direction and organization of Reasons and Instincts to make sure we can survive while achieving the Goal.

Knowledge Model

Formally, Knowledge is the Abstraction of Reality.

Our sensory data is organized in such a way that we have an abstract understanding of the reality we interact with.

However, the problem is the environment is always dynamic, and thus we need to be aware of the difference between the Reality and the Knowledge.

The discrepancy of the Reality and the Knowledge actually triggers the Risk.

Any difference between what happens and what we know carries a certain amount of risk, it is the risk which makes us uncomfortable.

Risk Model

GIGO (Garbage In, Garbage Out) is what we normally say in Computer Science, which implies that "Bad Input" will always produce "Bad Output".

GIGO is a defacto standard for Data Validation and Verification.

However, for any Intelligent Agent, GIGO is not enough.

Bad input can not only produce bad output, but it can also even harm the existence of the sentient agent.

Therefore, to be able to determine which input is good or bad is highly important even before actually taking the input.

We should be able to determine if the input is risky or not. It is the major function of the Instincts (Reactive Risk Assessment), in which Sentiments play the major role.

Limitations of Turing Machine

The real question of risk is how we could know the risk of the unknown input even before entering the input to the system?

Let me paraphrase this question: if some random keystrokes (inputs) can only not cause to display some errors from the computer, but it could also break down the computer, how should we design such computers to prevent this?

From this question alone, it is obvious to see the limitations of Turing Machine.

To decide the risk level of the unknown input even before entering the input to the system, some kind of mechanism is necessary.

It is where the Instincts come into play. The instincts are not as reliable as the reasons, but it is better to be safe than to be sorry.

The rule is we feel more comfortable with less risk. Happiness means less risk.

More than a Turing Machine

From this analysis, "Turing Machine" is not complete. It is not wrong, but incomplete.

For a Turing Machine to be able to compute "Risk Model", in addition to "Knowledge Model", additional components, other than Processor and Memory, would be necessary.

Allan Turing propose his Turing Machine with "Can a machine think?"

And now the question has become:

"Can a machine feel?"