Symbols of Power

by Sven Nilsen, 2021

What I am about to write is important to understand my view of Artificial General Intelligence (AGI)[1]. My hope is that by working through a key concept to AGI, I can later develop my thinking further by using this text as reference. This way I do not have to reintroduce the same building block.

One might think that a such building block related to AGI would require some technical definition of an abstract theoretical idea related to agent systems. What is surprising, is how close and central to the human condition[2] we come.

What is it that drives people mad? What were those countless wars in history about? What makes a person devote his or hers entire life to a cause? What makes governments pour billions into projects? Why do people fight for power?

All these questions are related to something I call "symbols of power".

A symbol of power is a system which has the following properties:

1. It is not well defined (undecidable[3])
2. It manipulates on a massive scale (powerful[4])
3. It has long reach across space and time (potent[5])
4. It never manifests physically directly (immaterial[6])

For example, gods, money, fashion, science and political power are not well defined, manipulates on a massive scale, has long reach across space and time and never manifests physically directly.

A scientist develops scientific theories and performs scientific experiments when doing science[7]. Although the scientific theories are written down somewhere and scientific experiments require some space, materials and work, the idea of science is immaterial. The scientist does not usually devote his or her life to some specific experiment or theory (this happens occationally). What all scientists have in common is the idea of science which they believe in. Governments pour billions into scientific projects, which to an outsider seems to produce no more results in the short term than a better math formula. In some experiments, several billions are spent just to check that the math formula works, that scientists already figured out in advance, sometimes decades beforehand. One such example is the search for Higgs boson[8], which required collaboration of scientists from many countries.

Is not this behaviour a bit peculiar and seemingly irrational? A scientist would say no, point to previous historical benefits of science and gratefully accept funding to do more research.

I chose science as an example above because it is a system which handles and welcomes criticism. The same thing could be done for other symbols of power, but it might not be as fruitful for discussion. Some symbols of power can not handle criticism at all. People being strongly influenced of these systems might experience anxiety and confusion when confronted with different views and opinions. This is why I define these 4 abstract properties (undecidable, powerful, potent and immaterial) such that one can reason about these systems without being worried about a concrete symbol of power.

Occationally, people might argue when or whether these 4 abstract properties apply to some system.

For example, a person believing in money might argue that money manifests itself in a physical form, although the physical form might vary from place to place. In some places people use coins for money, while other places people use computer algorithms.

What is important to understand is that the 4 abstract properties are simplifying assumptions.

A symbol of power might e.g. manifest itself directly at some rare occations, without this affecting the overall reasoning about the system. Thus, it is possible to apply the assumptions also in situations where people strongly disagree with making such assumptions in principle, perhaps because it would be offending. The point is not to offend people, but to produce a clearer framework of reasoning.

For example, Alice believes in god A and Bob believes in god B. Alice believes that at the end of time, god A will appear but not B. Bob on the other hand, believes god B will appear, but not A. Both Alice and Bob believe some god will appear at the end of time, but that the gods will not appear before that very last moment. Therefore, the assumption that the gods never manifest physically directly might be used about the time span before the end of time. Since Alice believes god B will never appear, it holds for god B for any time seen from her perspective, while for Bob the same is true about god A. This means that Alice and Bob almost agrees entirely and can apply the simplified assumption of an immaterial god in the foreseeable future, given that time does not end in the foreseeable future.

Since a symbol of power is immaterial, it does not have the ability to change the way physical objects change over time. Whether the symbol of power changes in a parallel time line is irrelevant. There is no way to confirm the change of these properties directly, because that would mean it could be observed as a change over time, like ordinary physical objects, and thus not being immaterial.

However, a symbol of power might change through the belief structure over time, because it is not well defined. There is no computer algorithm which people can write down that defines the symbol of power for all time, thus fixing its meaning once and for all eternity. Undecidability means that there exists no computer algorithm that can define a theory or solve some problem.

For example, money can change meaning in subtle ways over time, from being a system of exchange, to a system of global manipulation and environmental destruction through consumerism and inflation, or to a strange system people used in the past. It might happen that people in the future use a completely different system instead of money. We do not know yet and it is hard to imagine how, because we are used to think about the world through the perspective of money. However, we can not be confident in our belief that money will exists as long there are people in the future. A devoted free market enthusiast might believe that it is 100% certain money will exist in the future, because his or her friends are also free market enthusiasts and it is not cool to think it is only 99% certain. A person in the past, e.g. a trader delivering pepper and tea to Europe from India, might not have a clue what a free market is about, yet still use money in some sense.

People might also believe different things about the same symbol of power at the same time, while claiming to believe in the same thing. There is no requirement that a symbol of power influences people to have consistent beliefs. In fact, the property that a symbol is undecidable makes it very hard to confirm that two people do have the same beliefs. Most of the time, it can be that a meta-belief, believing that a group believes in the same thing, is sufficient to control the behaviour of the group.

A dictator is a person who claims supreme power over a state or a group of people.[9] Sometimes, a dictator uses storytelling and claims to be the physical manifestation of a god. This happened e.g. in ancient Egypt.[10] Notice that it is usually the dictator who claims to be a god, instead of the dictator claiming of becoming or replacing the god. A symbol of power can be so powerful that even a dictator does not dare to disrespect its authority, but instead "wears" the symbol. In terminology of Avatar Extensions[11], the dictator claims to be an avatar of the core of the symbol of power.

It is very common that symbols of power lead to introductions of beliefs in physical manifestations or miracles of the underlying "pure" form, the core of the symbol of power. One way to explain this is that the symbol of power manipulating brains, which are not stupid, realises that less potency follows from people understanding the symbol of power in abstract ways, recognizing it as a yet another tool for manipulation. The intrinsic conflict between immaterial form and manipulation of behaviour creates an innate feeling of being distant from or not belonging in the inner circles of the symbol of power, thus enforcing the desire to become like the attributes of the "pure" form. With other words, a belief of not being good enough, in some way, is often how the symbol of power creates the potential for control. However, if the distance is too great, then the symbol of power is less identifiable with the subject being controlled. This problem is solved by creating avatars which allow interaction or at least display that appeals to people's desire to form social relations, hence fighting for a place in the inner circles.

For example, in the fashion industry, models display clothing and how to wear them, which people then spend money to get the clothes and wear them to look like the model. Fashion in its "pure" form, or the core of fashion, is system that produces consumerism of clothing to maximise profit. The models are the avatars which humanises the core of fashion. It makes it easier for people to adapt a consumerism behaviour when they learn to know the models' names and background, admire them and desire to become models themselves. The industry selects people to be models, creating the inner circles which people desire social relations. A person wearing fashionable clothing becomes a local model, an avatar, in its surroundings, replicating the system and increasing the potency of the symbol of power.

It is not possible to write down a computer algorithm that defines fashion, yet it is a massive industry. Most people know what fashion means. Yet, fashion can mean different things to different people. To a businessman, fashion can mean wearing an expensive suit that changes in minor ways over time, requiring one to keep buying new expensive suits to stay fashionable, although the functionality of the suit must stay the same over time, which is to be comfortable to wear in long meetings.

One might think that fashion is being manifested physically all the time. Yet, if wearing a particular type of clothing was fashion, then one could simply keep wearing the same type of clothing. Fashion is the idea that clothing changes, it never stays the same, so fashion is unable to manifest itself as particular clothes. This is why fashion satisfies the immaterial assumption.

A symbol of power is not well defined, yet it is unchanging in an immaterial way and there are often very strict conditions that apply, which is part of the manipulation on mass scale. Furthermore, violating these strict conditions usually lead to penalties, so powerful that even a dictator might not dare to accent into the status of a god, but instead claims to be an avatar of an existing god. Avatars are indirect manifestations of the symbol of power, which "descends down from the heaven", or "has a place at the inner circle" of desirable social relations. Thus, it can be argued that symbols of power take on a life on their own (despite being essentially immaterial and unchanging in a physical sense) that is not depending on a particular person, except as an avatar of its core form.

The relation to AGI is the important property that AGI can create new symbols of power.

References

[1]   "Artificial general intelligence"
        Wikipedia
        https://en.wikipedia.org/wiki/Artificial_general_intelligence

[2]   "Human condition"
        Wikipedia
        https://en.wikipedia.org/wiki/Human_condition

[3]   "Undecidable problem"
        Wikipedia
        https://en.wikipedia.org/wiki/Undecidable_problem

[4]   "Power (social and political)"
        Wikipedia
        https://en.wikipedia.org/wiki/Power_(social_and_political)

[5]   "Potentiality and actuality"
        Wikipedia
        https://en.wikipedia.org/wiki/Potentiality_and_actuality

[6]   "Incorporeality"
        Wikipedia
        https://en.wikipedia.org/wiki/Incorporeality

[7]   "Science"
        Wikipedia
        https://en.wikipedia.org/wiki/Science

[8]   "Higgs boson"
        Wikipedia
        https://en.wikipedia.org/wiki/Higgs_boson

[9]   "Dictatorship"
        Wikipedia
        https://en.wikipedia.org/wiki/Dictatorship

[10]   "Ancient Egypt"
        Wikipedia
        https://en.wikipedia.org/wiki/Ancient_Egypt

[11]   "Avatar Extensions"
        AdvancedResearch – Summary page on Avatar Extensions
        https://advancedresearch.github.io/avatar-extensions/summary.html