What we can learn from unanswered letters to extraterrestrials

On the aesthetics of cognition and communication

Georg Trogemann
Academy of Media Arts Cologne
May 2020, Barcelona

I am sitting at a table in the living room of my flat on the 6th floor of a building in Barcelona. From time to time I hear the gas sellers beating their gas canisters with iron sticks. For me that sound is a piece of information. I know that I could now walk onto the terrace, give a sign and the seller would come to sell me a canister of gas. Since I don't need gas today, I stay seated and continue reading at the computer. The decision not to react, to stay seated and continue reading, is nevertheless a decision I make.

Surely nobody would deny that this situation requires **cognition** or what we call today **information processing**. By cognition I mean "the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses."[1] A person's cognitive abilities include perception and attention, memory and learning, knowledge and emotions, problem-solving, creativity and imagination, concept formation, recognition, planning and orientation, language, deductive reasoning and decision making, introspection and self-description, will and faith. Obviously not all these characteristics are necessary to speak about a cognitive system. There are much simpler organisms that we nevertheless accept as being cognitive. In fact, we must be aware that the cognitive sciences are a huge interdisciplinary field comprising philosophy, psychology, anthropology, neurosciences, artificial intelligence and linguistics. In other words, a field in which one can fail wonderfully. So, the issue of the intended work needs to be focused and narrowed. My questions are: If you encounter an entity (let's call it system), how would you decide that it is a **cognitive entity**? Or perhaps better, how can you find out what level of cognition and intelligence it achieves? At what point does it make sense to try to **communicate** with this entity, what could the communication look like? Do these questions—like thinking in levels of cognitive capabilities—make sense at all, or are they misplaced and just the senseless production of a thinking mind?

Some words about the strange and somehow contradictory subtitle. In philosophy, **aesthetics** means either the theory of sensual perception, or a theory of art or design. In a direct reading, the aesthetics of cognition would therefore be about knowledge and understanding from the

[1] "Cognition" in Lexico. Oxford University Press and Dictionary.

perspective of its sensual perception, or even more directly about the multiple forms of sensually depicting thought. But we can describe it an easier way. With the term "aesthetic" we actually refer to a set of notions and concepts that underly particular works and theories about cognition and communication. So, with the aesthetics of cognition and communication we address general principles of cognitive entities, be it organisms or machines. The goal is to play around with the basic ingredients of cognition and communication experimentally and open them up for speculation and the unrealized possibilities of cognitive systems. On the basis of experimentation and the new forms of depicting it, symbolic media, AI programming, mathematics and natural languages can be examined as extensions of the mind.

To deal with the situation of the gas sellers in Barcelona in the appropriate way, one needs what we call a representation of some part of the world. I have often heard the clang of the sticks against the gas canisters. I can recall pictures of the gas sellers in my memory, how they react to hand signals from balconies, and deliver their canister to buyers. Someone who wishes to buy a gas canister must have an idea how much gas they have left and how long it will last them. While I am sitting at the table and listening to the metallic sound of the gas canister, I cannot see the gas seller. Nor can I see the gas canisters in the storage room and I have no access to how full they are. So the right decision at this very moment depends on conditions in the world to which I have no access. In fact, the sound that comes to my ear from the street is only information because it meets structures in my head that somehow represent a part of the world and which have meaning to me. At the same time this decision could easily be taken by an information processing machine. The necessary parameters can be stored and processed considering various other values like price, expected future usage of gas, etc. However, this would also be a very crude and emblematic realization of the problem. Representation can be much more sophisticated. But at the same time the action could not be done by a slime mold. Nevertheless, slime molds are also cognitive systems, as we have learnt from biophysics.

"Most of what matters, most of what an agent cares about—everything in the past, everything in the future, everything that is far away, and a huge number of facts about what is close by, too—cannot be directly detected. […] Yet even if we cannot directly detect it, we need to know—if we are intelligent—what is going on, distally: where we stored those acorns, that a predator lurks behind that rock, that the person who just left the room still exists, rather than having metaphysically vanished, that the sun will rise again tomorrow. […] How do we orient appropriately to a world beyond effective reach? **By making use of locally effective structures and processes in the world's stead—that is, by representing the world.** This is essentially a metaphysical requirement on intelligence. [Those] representations have two criterial properties: (i) they stand in noneffective semantic relations with what is distal, with what is not immediately effectively available (what they represent); and (ii) they can nevertheless effect or be affected by physical changes within causal proximity."[2]

According to Brian Cantwell Smith, any intelligent system must fulfil some crucial principles connected to representation. His insight rests on fundamental theses about representation and the nature of physical causality that systems have to comply to if they allow "reason" or "process information".[3] The system works or functions (mechanically, physically, biologically) in ways explicable by science on the basis of cause and effect. But the system's behaviour and constituents support and need semantic interpretation—that is, can be taken to be about (mean, represent, etc.) facts and situations in the outside world. Semantic interpretation here means that we need the outside world and its meaning to understand what the system does. It is—for example—not sufficient to look at the neurons in the brain and their patterns to explain what is going on. If asked "What do you think when you hear the sound from the street?" I would not answer in terms of

[2] Smith, Brian Cantwell. *The Promise of Artificial Intelligence* (The MIT Press) (p 14-16). The MIT Press. Kindle-Version.
[3] Ibid., p 9.

neurons and electrochemical processes in my brain, but refer to things in the outside world, about gas sellers, levels in gas canisters and the frequency of my cooking. "What matters about such systems—whether they are right or wrong, true or false, useful or useless—has to do with an assessment of what they are interpreted as representing or saying."[4] This means we have to be aware that the interpretation as well as the assessment of cognitive systems takes place outside. In cases where we build such a system ourselves, we assess them by the functions they provide anyway. In the case of natural systems, we attribute the evaluation and objectives of the system from outside. They cannot reliably be measured at the system itself; only its behaviour can be interpreted. "In general, semantic relations to the world (including reference) are not effective. The presence of a semantic relation cannot be causally detected at either end—neither at the signifying object or event (term, word, representation, data structure, thought, etc.) nor at the signified entity (reference, denotation, or represented entity or state of affairs)."[5] In simple terms, you cannot measure whether somebody thinks about the gas canister from the gas canister, nor at the brain that a certain state is representing the gas canister. That is actually one of the first questions: what does it mean to be a representation of something? It does not mean, obviously, that the representation is in any sense similar to what it represents. It also does not mean that it has to be locally restricted and static. In how many different ways can a gas canister be represented? In this context, external memorization is interesting, as is realized by ant pheromones that allow ants to find the shortest way between a feeding place and an anthill. **It is important not to see representation like a map of some reality or a detailed model!** But something about the world must be stored that, together with the outside circumstances, allows the desired behaviour to be produced.

Cause-and-effect mechanisms in rational, information-processing systems are therefore interrupted in two directions. Firstly, one cannot tell from observing and measuring the things in the world that they are objects of a representational process. Conversely, one cannot easily determine within the information processing system (because of the multiple realizability of any representation or information) which semantic information is represented. This makes it very difficult to decide whether a supposedly intelligent system (for example a slime mold) is a cognitive information processing system or not. In order to achieve this, it is necessary to decipher the coding of information and the information processes going on within the cognitive system. Whether information processing is a sufficient criteria for intelligence is a question I am not interested in. Today we would say "no", fifty years ago we might have answered "yes". But representational information processing it definitely sufficient to be able to speak about systems that can make decisions, to look into the future and to maintain a space of different behaviours within a world that is for the large part beyond reach.

So the question is, although causality is interrupted in terms of thinking in both directions, can we still say something about cognitive processes and under what conditions we can even communicate? The only option is to observe these systems at work, that is, in their dynamic behaviour. Under what circumstances is it possible to decide that the system uses representations of the world, that it processes information or is able to learn? What sort of convincing and defendable explanations are possible, how do they differ from how the system (white box) really functions. In our example of the gas sellers in Barcelona we can derive certain capabilities from observing the situation. If I do not react to the sound from the street, one can of course conclude nothing. But if I react and stand up to buy a canister of gas, we can conclude from that behaviour that representation and information processing must be involved. We can, to a certain degree, measure it by observing the levels of the gas canisters in the storage room, to which I did not have direct access at the moment of decision making. The goal is, of course, to understand how this performance is achieved.

[4] Ibid., p 11.
[5] Ibid., p 12.

It is clear that it is almost impossible to say anything about the cognitive ability of an entity as long as the entity only perceives and does not act itself. If there is no action, you cannot really draw conclusions. As we know: "If you had kept silence, you would have stayed a philosopher."[6] Systems give themselves away by responding. So we are especially interested in thinking processes that express themselves in actions. This is not simple behaviourism, because we want to include the inner structure of the entity. Of course, the matter is most simple if the structure is known, as in the case of technical artifacts. In other cases, it may be possible to make measurements that allow conclusions to be drawn, or it may be possible to draw conclusions about the minimum complexity of the internal organisation from the complexity of its actions. As long as the cognitive system controls physical cause and effect chains, we can hope to find out something about its internal organisational complexity. They make use of locally effective structures and processes to reach certain "goals", they are able to control their physical environment.

The starting situation for my simple (thought) experiments is therefore the following: I assume that cognitive systems are material realizations (physical, biological, etc.) that obey physical laws. But to be able to speak of a cognitive system, as Brian Cantwell Smith convincingly argues, physical causal chains have to be reorganized. Our physical universe is a network of local influences. "It is a consequence of the most basic fact in all of physics: that causal influences fall away locally in both space and time. The fact that the universe is a web of purely local influence poses a formidable challenge to any system or creature aiming to comprehend its world."[7] To overcome this restriction of local causality and to cope with a world out of reach, cognitive systems must manage to represent parts of the world and act with those representations instead. **So, the basic level from which an examination of cognition should start is this level of representation, which is already beyond pure physical causation.** Any cognitive entity is constructed on this primary level. It is this level and not physics, were we must settle all knowledge about the world. For knowledge itself is a product of cognitive systems. What we know about physics, i.e. what we think physics is, must first of all be distinguished from the physics that realizes the cognitive systems themselves.[8] In fact, the study of cognitive systems starts with indissoluble dualism. The dualism of subjective perception and objective description. The effect of the world expressed in cognitive systems versus the models of the world that the cognitive systems produce of themselves. I would like to start with minimal systems that represent a part of the world and then go further to **memorization**, **simple processing**, **learning**, **possible spaces of action**, **communication**, and the **distribution of cognitive processes**. What I don't want to consider here are concepts such as consciousness, affect, emotion, reflection or introspection. The goal is not to find new universal mechanisms (like the Smith metaphysical requirement of representation), but to use the ingredients of cognitive systems and to play with them and come up with some experiments that allow us to reflect on cognition. In this sense it is an aesthetic approach. Even if it should become apparent in the future that sensitivity to pain and affect are the basic prerequisite on which all higher cognitive performance must be based, it is obviously possible to build certain cognitive systems that can do without it completely. Affect and emotion are two terms that are subject to what we call the initial dualism of human cognition, they are (internally) subjectively experienceable and (externally) physically describable. These Janus faces are the fundamental initial conflict of human cognition caused by the basic representational configuration.

---

[6] Boethius, *Tröstung der Philosophie*. "Hättest Du geschwiegen, wärest du ein Philosoph geblieben."

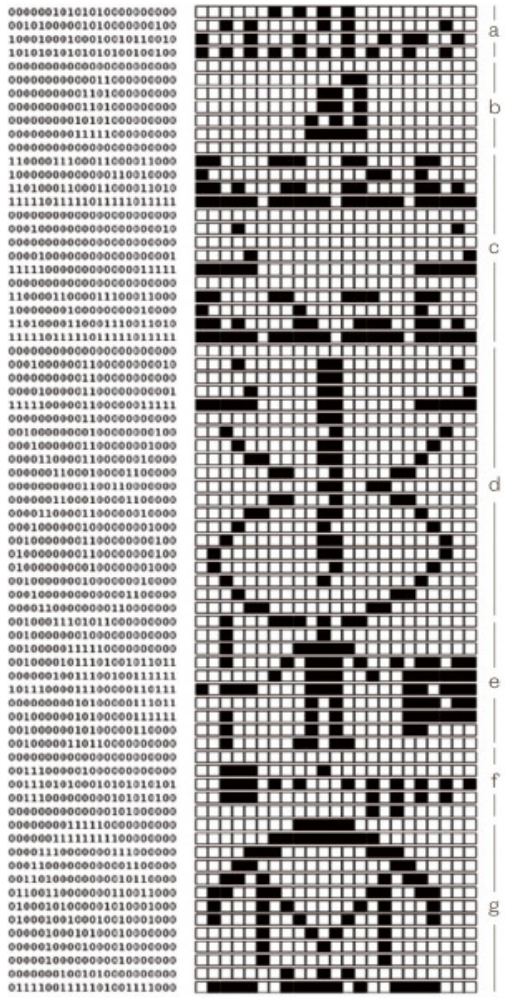
[7] Smith, Brian Cantwell. *Artificial Intelligence* (p.14).

[8] So we must start from their differences and seek convincing arguments if we wish to bring the physical, and our model of it, into congruence. Again, such congruence is only the result of a cognitive process, which can be supported by measurements in the real world. In my opinion, however, we should, in the sense of Richard Rorty, give up on finding an inner mirror of outside nature from the outset. Evolutionary epistemology thus argues with the fit between reality and our idea of it.

Some fields where we can find inspiration for experiments:

1) **Representation, plant intelligence and animal communication**

Closer examination of plants and simple organisms makes it clear that the question of cognition is above all a conceptual one and that cognitive entities cannot be easily distinguished from non-cognitive. "Potential cooperative behaviour between plant species has been observed in which pairs of species seem to prefer to grow adjacent to each other. The use of volatiles in cooperatively helping protect neighbours against herbivore attack has been reported. Competition for space between trees occurs particularly with close neighbours and leads to crown asymmetry. The strong indications are that crown asymmetry is adaptive behaviour and thus can be regarded as intelligent behaviour. Root systems in a number of species have been shown to have the unusual properties of self-recognition. It is thought that self-recognition operates through physiological coordination. Recognizing one's self is of value in that, with a plastically developing plant, self-competition among growing roots is reduced and internal resources saved."[9] Is adaptation intelligence or do we have to tell them apart? On the other hand, what we could call a long-term memory of trauma obviously does exist in plants.[10] Another field of inspiration for experiments would be the communication between animals.



Arecibo

2) **Communication with aliens, letters to extraterrestrials**

"The search for extraterrestrial intelligence (SETI) is a collective term for scientific searches for intelligent extraterrestrial live, for example, monitoring electromagnetic radiation for signs of transmission from civilizations on other planets."[11] A message (the so-called Arecibo Message) was sent from Earth to possible extraterrestrials in the form of a radio-wave signal. It was activated only once on November 16, 1974 from Arecibo observatory, the second largest radio telescope in the world, near Arecibo in Puerto Rico. The message consists of a total of 1,679 bits. A receiver of this number of bits should decompose it into its prime factors 23 and 73 and then split the sequence of bits into a 23×73 matrix as black and white image. The message contains binary coded information about the biology of the human being as well as about the human population and the origin of the signal. The message is written very anthropocentrically and it is very doubtful as to whether a non-human extraterrestrial could do anything with it. In fact, with this message we have delegated problems of philosophy and metamathematics, for example whether arithmetic is based on analytical or synthetic judgements (Kant vs. Frege), to extraterrestrials. The first thing we would expect from an extraterrestrial intelligence is that they know how decomposition into prime factors works. Nevertheless, from this field of research we can learn a lot about the problems of identifying

[9] Anthony Trewavas, *Plant Behaviour and Intelligence*, Oxford University Press, 2014, p 181.
[10] Daniel Chamovitz, *What a Plant Knows – A field guide to the senses*, Scientific American, New York.
[11] SETI – Search for Extraterrestrial Intelligence, Wikipedia, visited on 8 May, 2020.

cognitive entities and starting basic communication with them. "Given the level of abstraction of the concepts used in systems theory, it is no coincidence that a fundamental concept of Niklas Luhmann's theory of society describes perfectly the encounter between humans and extraterrestrial intelligence, namely, the concept of double contingency. This refers to the situation where two systems that are completely unknown to each other perceive each other and are aware of this situation; in other words, they perceive that the other system is perceiving them and are aware that this is so. The crucial factor is that both systems are non-transparent to each other. They are black boxes: only input and output can be observed; the internal, self-referential operations of each are hidden to the other system. This situation of double contingency is a fundamental theoretical figure in Luhmann's theory of society. It is a situation of double indeterminacy of expectations and expectations of expectations as well as of assumptions and of assumptions about assumptions—in short: it is a situation of the greatest possible mutual mistrust. It is the point of origin for every evolution of social systems."[12] The systems theory view is very close to my abstract cybernetic approach. Another source for an introduction to the problem of identifying intelligent beings and communicating with them is the NASA report by Douglas A. Vakoch.[13]

3) Possible action spaces

In order to be able to speak of a cognitive system, the entity must have a scope of action. Even if we are only dealing with a binary decision, there are two options from which to choose, and this choice must not be random, but should be "freely selectable" (whatever that means, since an algorithm has no real choice, but must follow its rules) and "better" in some sense than the other with respect to a particular situation. Spaces of action, even if they cannot be explicitly named, are related to representation and are also a metaphysical necessity in order to speak of cognition. In that respect, every language offers more possibilities to express than can be used by the individual in their lifetime. Even the whole of humanity cannot exploit the potential of a language. But in what sense do these sentences exist, if they are never spoken nor heard. The space of possibilities of language is characterized by a surplus compared to the reality. For the sake of its representational capacity, it must always have more possibilities than reality exploits. In the same sense, images form a cognitive space of possibility and in this sense AI algorithms are methods for establishing certain meaningful connections between images in this possibility space. In organisms and living beings, spaces of possibility are realised, for example, by genes or neuronal nervous systems.

4) Embodied mind, enactivism, situated cognition, extended cognition, social and distributed cognition

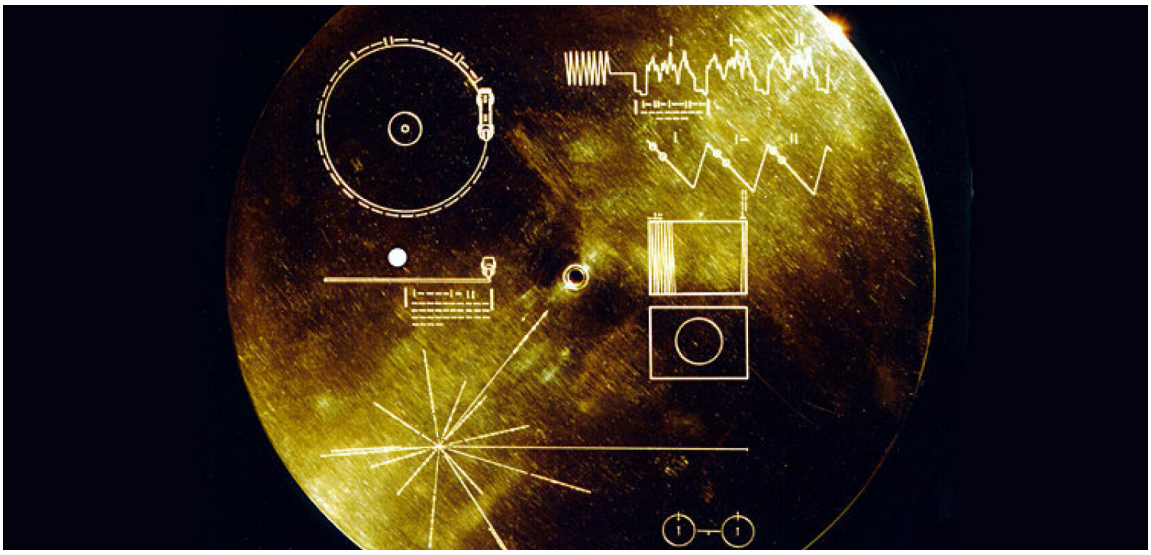
Leroi-Gourhan's book *Gesture and Speech* took an integrated approach to human evolution. He conceives gesture as material action, and as indivisible from speech. They are products of an embodied mind that spawned our technical and social achievements. His reading of the evolutionary process between the hand and the face provides a biological basis for cognitive as well as communicative aspects of our culture. There are of course other approaches that reject the representation paradigm. For example, enactivism is critically distinct from the cognitivist paradigm, in which cognition is seen as the manipulation of internal representations of an environment independent of the system. Enactivism is based on the assumption that mental processes are dynamically generated by the embodied interaction of a cognitive system with its environment. But this is not necessarily in contradiction with representation. It is more a question of what a representation is. No one has yet been able to explain how certain brain functions should arise without a form of memory, and memory is representation. A deeper look at the different

[12] Bernd Pröschold, *Communication with extraterrestrial intelligence. Observations based on systems theory*, translated from the German by Kathleen Cross. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3294358

[13] *Archaeology, anthropology, and interstellar communication*, edited by Douglas A. Vakoch. https://www.nasa.gov/sites/default/files/files/Archaeology_Anthropology_and_Interstellar_Communication_TAGGED.pdf

architectures of cognition and the structure of cognitive systems might provide ideas for tiny experiments.

When we ask ourselves what kind of freedom entities (humans, organisms, animals, machines, etc.) can achieve in relation to their environment, there are two interesting strategies. Firstly, we can work with realizations that we find in our world and deliberately manipulate them to find out how they function (as is done in animal experiments for example) or secondly, we can start to build those entities ourselves. I am interested in the latter, i.e. purely technical systems and accompanying thought experiments. If we succeed in realizing systems that we would undoubtedly, and from all points of view, call free entities, then they are free, even if they are built by us. It actually wouldn't even matter if we ourselves were deterministic machines. Even if we are prisoners of determinism, for example the deterministic simulation of a matrix-like system, as long as we have the freedom to take sub-captives that we can free, we are somehow the masters of freedom.[14] In this way we prove our own freedom. I therefore suggest constructing a series of speculative experiments for so-called "cognitive systems". The plan is to **build (or simulate) and speculate about cognitive systems** that are obviously able to represent parts of the world, to process information, to learn, to adapt and evolve over time. Since we will have constructed these systems, we will know how they function. Then we forget about that knowledge and try to explain these systems by observing them and taking measurements. In this way we somehow use white boxes and treat them as black boxes to see what we can find out. This is **black box/white box gaming** about metacognition and about communication. Since I am not primarily interested in what those systems do, but how they function and what they can achieve in principle, we are in the field of metacognition. Informatics normally wishes to build systems that provide a certain performance or solve an ambitious problem. I want to examine very tiny systems, that are probably uninteresting from the performance perspective, for theoretic reasons and as philosophical gadgets and means of thinking.



Voyager golden

[14] This was an idea of Otto Rössler regarding Werner Fassbinder's 1973 film *Welt am Draht*.