Strong Semantic Computing, a BICA framework

Abstract:

Standard computing will be characterized as a functional extension of syntax. This is based, partly, on Searle's *Chinese Room* argument. Using BICA philosophy, in particular the claim of continuity between human-animal-robot cognitive architectures, we will define strong semantics as the ability of a cognitive architectures to consult cognitive maps, in particular *phenomenal content* map. The goal of strong semantic computing in autonomous robotics should be to 'know what is going on' before engaging in detailed logical analysis (it can be called *Gestalt computing*). Such computing is needed in advanced autonomous robotics, especially robots functioning in human environments. Incidentally, this approach provides a partial solution to Searle's *Chinese room* case.

1. Unity of cognitive architectures
    a. BICA as a continuum of human, animal and artificial cognitive architectures

The early message of the BICA movement may have been that AI can learn from neuroscience. I still remember discussion among board members of *The International Journal of Machine Consciousness* (in 2009), when only Bernard Baars and I defended this claim (A. Samsonovich joined the board later), and prevailing view of AI experts was still a negative. Some eight years later the issue is no longer contested, seriously. The current message of BICA is that there is continuity between human, animal and artificial cognitive architectures. The main challenges to this view are: 1. View of human language as a qualitative step over animal cognitive architectures. 2. View of biologically based consciousness, some kind of *mentations*, as an essential, long term, qualitative step over AI. Challenge 1. prominent in mid-20th century philosophy of language (most notably by D. Davidson's dismissal of thinking animals), denies smooth transition between humans and other animals. Step 2., raised by philosopher-linguists (most notably J. Searle's denial of 'original' intentionality outside of biological organisms), denies smooth transition from animal cognitive architectures to AI. This last step relies on a biological essentialism in defining intentionality through *the mental* (Brentano). The problem here is not the last clause, but the claim that 'the mental' is not, and under normal circumstances would not be, available to artificial cognitive architectures. This claim has been the target of my work on *the engineering thesis in machine consciousness* for the last ten years. In this paper I argue the following: Homo-centric notion of intentionality is the gist of the very definition of semantics. Hence, the prevalent old-school notion of semantics is a major problem in preserving the smooth transitions (humans-animals-artificial agents) in BICA. Even new-school notions of semantics, in cognitive linguistics, do not leave the old conceptual framework radically enough. Naturalistic notion of semantics, and rejection of the semantic view of intentionality (Chisholm), allows us to avoid Quine's dilemma (having to choose between intentionality and physicalism).

Practical implications of the unity of cognitive architectures approach allows:

i. Doing away with constraints on cross-engineering (engineering sciences)
ii. Doing away with constraints on cross-engineering (humanities and law)
    Here is a more detailed explanation:

b. Weak semantics

Weak definition of semantics: 'Semantics concerns the meanings of words, signs, symbols, and the phrases that represent them.' We know, mostly from the context, in what sense an expression, e.g. that an X is 'a piece of cake' has been used (literal, figurative etc.) https://www.alleydog.com/glossary/definition.php?term=Semantics . Yet, hints how to use terms of common language come from advanced observation of syntactic regularities, as well as from a broader context of use, pragmatics. (Those hints can be opaque but this is not specific to thinking machines; even competent human users may disagree on the proper meaning). Within what we call the weak definition of semantic computing – neural networks place words in semantic positions new to them based on those syntactic (or, contextual) observations. In this framework, semantics is just a short extension of syntax (often within pragmatic context).

c. The Sanz's challenge.

I first met Riccardo Sanz, one of the European leaders in autonomous computing, at a conference in philosophy of mind; *his interest in the topic was a bit puzzling*. When I asked him about the reason for his interest, Sanz explained that adding functionalities to functionalities, e.g. in building a self-driving car, seems to meet some limits. While it works in delimited environments, it does not work in complex real-life situations. Even an individual with rather low IQ understands how to comport oneself in a complex human environment, in which a very smart robot makes massive mistakes. In order to avoid those mistakes an AI needs 'to know what's going on'; the best way of doing so is not by adding functionalities.

Knowing what is going on is what we shall call the Sanz challenge for AI. Weak semantics as defined above is unlikely to meet the challenge. Cognitive architectures need to 'see' the big picture before they start 'thinking' in terms of semantics and propositional logic.

d. Beyond myopic computing

Try to have a close look at a complex image, e.g. a realistic painting of a river at a sunset. It does not strike most people as a very difficult thing to grasp. Yet, try to describe it in language. It is easy to give a general description of the gist of the theme and composition of the painting. However, try to give a one-one description of the painting, on the basis of which one could reproduce such painting with at least 90% of *truth preserving.* Oh well, this would be a very complex, long description indeed. We need to avoid the need for such translation in semantic computing.

2. Semantic computing – human level phenomenal maps

a. PHENOMENAL MAPS *as the basis of human-level semantics*

Semantic means of communication are clumsy means to transmit phenomenal information. We seem not to think primarily in language (contra Davidson); neuroscience makes a strong case that we think using complex, overlapping maps (Damasio) – those are perceptual, activity-based, or more abstract (categories like higher/lower or cold/warm available to non-human animals). They help us address enquiries, for instance respond to the question how many kayaks are on the picture. Keeping all the answers to possible questions in a database is a waste of memory – one needs to be able to consult a map instead.

Such maps can be external (like a picture on a wall, or a view out of the window), internal (memory based), kinetic (based on muscle memory activated in interactions) and so on. As visible in neural nets, much intellectual formation comes through interactions, the ability to perform a task, to improvise, or to hold a conversation.

Mental maps are set up, and constantly verified, in an unending dialectical interaction among:

i. External pragmatics – influences *objective* ontology (consistent with externalist theory of meaning}
ii. Internal pragmatics – phenomenal-functionalist *qualia* (consistent with internalist model of meaning)

At the final level of analysis we encounter retranslation to broadly phenomenal maps:

iii. Reduction of external to internal pragmatics (contra O'Regan)

The level where mental maps overlay consciousness is phenomenal experience. Whatever passive or active experiences took place they all come, through the nervous system, to CNR to be put on consciousness – unless those experiences are unconscious. The ontological status of whatever phenomena may represent is an inductive hypothesis, always subject to further investigation of status; phenomenal *gestalts* are included (we are not advocating a *sense-data theory*). Their meanings, based largely on functions, are confirmed in an always unfinished, never definite set of inductive hypotheses with various levels of verification – and those verifications shape up in the framework of broader conceptual frames (paradigms).

b. Phenomenal markers translation of phenomenal and functional levels of description (Baars)

Why do we need phenomenality in the experience? B. Baars claims that it is one more way to discover the relevant features of the world – just like colored folders, in most situations, make it easier to pick the right one then the ones with monochromatic covers. Humans detect specific structures of the world by color, smell, touch, temperature, taste etc. while some animals have features that are dormant or weak in humans (such as echolocation). We shall not claim that humans and other animals exhaust the set of potential phenomenality – this opens the topic of extra-strong semantics.

3. Extra-strong semantic computing

The thesis of this section is that we should expect machines to develop their own and in many cases improved semantics based on more advanced phenomenality than that humans and animals. While this topic seems unpopular among most contemporary philosophers, Spinoza's speculation about angels opens interesting possibilities.

a. Machine to machine analysis

First, let us analyze situations where machines seem superior to human beings not just in terms of brute computational force but by having specific, and probably superior to human, modes of doing things.

i. Computer proofs as superior to human proofs

Many theoreticians are puzzled by computer-generated mathematical and logical proofs. Some of those proofs may be translated into alternative mathematical notations, which allows them to hold intuitive value to humans. We can call them Kauffman translations.

Yet, many machine proofs are not translatable into proofs that are intuitive to humans. Some authors believe that these are no proofs, just some pragmatic demonstrations, unless a mathematician can understand them (Hadley), while others find them somehow inferior since human beings are unable to learn from them (Copeland). I am tempted to view purely machine-readable proofs of mathematical theses (so far) unprovable for humans as superior to human-provable theorems – this is because they show radically different, thus novel methods of 'thinking'. If human beings cannot learn those means of thinking, and communication, in our present biological forms – it may be possible to use moderate enhancements to attain this task (just like our eyes can see infrared with special googles), but in some cases even this may be impossible.

ii. Spinoza level maps

If there is no-human translation, we have what I shall call a *Spinozian cognitive architecture*, which is non-human accessible. According to B. Spinoza there were just two categories that human beings could use – let us call them phenomenal perceptions (related to the sense of the self) and object-based grasp of things (that takes for granted ontology the way physicalist reists do). B. Russell developed this though in the 1922 version of neutral monism. Yet, Spinoza claimed that there are infinitely many different modes of presentation of the world, which are available to angels. In this conceptual framework, epistemologically, artificial cognitive architectures may be approaching Spinozian angels. Using the phenomenal markets view towards machines, a machine could, for instance, 'perceive' different spin of elementary particles as a relevant carrier of information, whereas human perceptual sphere is much lower.

4. Conclusions
    a. Weak semantic computing based on filling in the blanks or posing specific questions does not help machines to learn 'what is going on'. It is overly myopic and step-by-step.
    b. Strong semantic computing based on perceptual maps, always leveled to broadly 'phenomenal' maps, allows humans to acquire the general picture view. Such view allows us to look out of the window and see a landscape not through complex syntactic structures but through general images; such images accompany superior thinking in the games of strategy (for players of comparable computational power), *and even mathematicians who tend to visualize when building a novel proof*.
    c. Extra strong semantic computing based on Spinozian maps, allows for broader than human perceptual and information processing level in future AI.