Systems and Consciousness

Ricardo Sanz, Juan Escasany and Ignacio López

Autonomous Systems Laboratory

Universidad Politécnica de Madrid,
José Gutiérrez Abascal 2,
28006 Madrid, Spain
ricardo.sanz@etsii.upm.es

Abstract: Consciousness can be explained. This paper proposes a view of consciousness that fits biological and engineering perspectives. We answer some basic questions about consciousness from this perspective and conclude with some proposals for further research from the perspective of intelligent control.

Keywords: Consciousness, Intelligent control, robotics, autonomous systems, intelligence, physically grounded systems, constructiveness.

I. Introduction

This paper is a contribution to the conference that tries to provide sound answers to some basic questions that pervade our disciplines

Issues that will be addressed in the conference include all classical topics about consciousness that were manifest in the questions posed in the call for papers:

- Is consciousness a physical process?
- What are the neural correlates of consciousness?
- How can we build a systematic theory of consciousness?
- Is consciousness connected to fundamental physics?
- What is the relationship between the subjective and the objective?
- To what degree are animals conscious?
- Can machines be conscious?
- How does consciousness fit into the social and cultural order?
- Are naturalistic accounts of consciousness circular, given that experience of nature is a construct of the mind?

This paper will provide answers for all of them from a control systems perspective that we think is the proper stance to talk about minds.

II. Bodies, minds and control

The need for bodies to have minds is not new at all. It is not a discovery of the situated robotics movement. In 1831 William Godwin, in his thoughts about man, said [**?**, pp.18]:

"We have every reason to believe that the mind cannot subsist without the body;"

Intelligence is manifest in the exploitation of information to perform better. This view is strongly related with the classical interpretation of rationality [**?**].

We would expect that any symposium on Intelligent Control should be focused on the "intelligent control" topic, and with this idea in mind we have been coming to these events during the last years, to learn something about it. Unfortunately, nobody here seems to know what does "intelligent control" mean. We have found people that know everything about expert control, or neural nets, or the use of fuzzy technology in control systems. But nobody knows about intelligent control itself.

We all have failed in the elaboration of a common understanding of the term; we have failed in the generation of a shared meaning that could be used to serve as a basis for sound research an development. We discuss from our personal stances without reaching a meaningful agreement. Some years ago, the IEEE task force concluded that everything would fit under the umbrella if anybody would like to do so.

III. A Theory of Intelligence

In the NASA Highly Autonomous Systems Workshop of 1997, Marvin Minsky "suggested that a theory of which AI techniques actually work, of how well and in which domains, was probably achievable at this point" [**?**]. Why are we striving for local achievements when we should be working in a global theory of intelligence?.

We need this theory because the type of work we are doing now (*Guys, look at this neural net! See how well it performs in this very specific problem!*) does not work to make engineering at large. And engineering-at-large is what we need to build highly autonomous systems like those needed in deep-space applications, intensive medical care, risky industries, airplanes or future military systems.

Maximum autonomy depends on the level of consciousness achievable

Topics related with this research:

The model of the self in the model of the world

IV. The nature of Consciousness

What is *consciousness*? What is the meaning of the word *conscious*?

Conscious just means "ON". It can only be applied to a semantically closed systems [**?**]. A human is a semantically closed system. A pseudomonas aeruginosa is a semantically closed system. A robotic control system is a semantically closed system. When they are ON, they are conscious.

In particular, consciousness refers specifically to the

ON/OFF state of subsystems for information capture of a semantically closed system. Conscious means that the system is perceiving, *i.e.* generating mental representations of the world —including itself— that can be used to act better.

Each perceptual mechanism can be ON or OFF. This means that in multiperceptual systems there are several consciousness states. *Altered states* in humans are just consciousness states different from the normal state (some OFFs where there were ONs and some ONs where there were OFFs). IMHO, human altered states are always ON-¿OFF states. Psychedelic drugs were thought to provide transitions to altered OFF-¿ON states *i.e.* states where we can perceive *more things*. But the perceptual mechanisms they open are in most cases are just internal mechanisms (signals from the brain).

In general, altered states modify priorities of perceptual tasks (the ON-¿OFF is just a extreme case with binary priorities) being directly correlated with mechanisms of attention. Perceptual policies like this have been applied in many control systems. [**?**] describes the use of expert system focalization mechanisms to focus control system attention in a specific, high-priority, task.

Top level self-consciousness —perhaps the central topic of the conference, is just the operative condition of the perception subsystem that generates the mental representation of the self inside the mental representation of the world.

In this sense, most machines are not self-conscious, because they do not have representations of themselves, but it is not the general case. Situated robotics, functional reasoning systems, fault-tolerant controllers, Turing-test challengers, *etc.* keep information about themselves that help them achieve better levels of autonomy in uncertain conditions.

V. The answers to the questions

A. Is consciousness a physical process?

Yes. Everything we can talk about with sense is a physical process. Metaphysical processes could —in theory— exist, but they are not relevant for humans because they do not interact with us or with our world.

In my personal opinion, metaphysical processes do not exist. Their existence is restricted to spoken, printed —and sometimes engraved— words. Is the same case of fairies or Loch Ness monsters.

B. What are the neural correlates of consciousness?

There are no neural correlates of consciousness. Consciousness, as commonly perceived, is a global property of a system. It is not a particular state of it or of a subsystem of it. Searching for consciousness correlates in the brain is like searching for *vehicleness* correlates in a car or for stability correlates in a dynamical system.

They are behavioral descriptions refering to the interaction of a system with an environment. Consciousness is a property of systems that exploit information interaction with their environment, with their world [**?**].

C. How can we build a systematic theory of consciousness?

We already have that theory. It is the theory of *signals and systems*, or more concretely the theory of *systems*.

The main problem we have is the degree of difficulty we find in extending the analytical capability of this discipline to systems of higher complexity.

In systems science and engineering, we like sound mathematical models (science) and sound methods to exploit these models to achieve objectives (engineering). Complexity poses an effective barrier to this endeavor, because we do not have good mathematical tools or abilities to deal with systems that are: chaotic, non-linear, time variant, large, highly coupled, *etc.*

In these situation we have two main alternatives: reduce model complexity maintaining soundness of reduce soundness maintaining model complexity.

An example of the schism we are experiencing inside the area is the everlasting discussion about the real relevance for our *scientific* community of soft-control theories like fuzzy, neural or genetic control.

People from the *human* sciences do not suffer such a schizophrenia —at least not as big— because they have been always dealing with such complex systems as humans, and their theories have always been soft theories.

But it must be clear that softness is undesirable for several reasons. I will mention two:

- Soft theories aren't good because they lack predictive power. They don't give us high degrees of certainty in relation with the evolution of the world or the effects of our actions (think about economy, sociology or even some parts of medicine).
- Soft theories induce the metaphysical way of thinking carrying us to theoretical places were we shouldn't go as scientists, because it is *so difficult to distinguish a soft theory from a meta theory.*

D. Is consciousness connected to fundamental physics?

I understand that the question can be interpreted in three different ways.

If it is asking if we need quantum mechanics to *explain* consciousness I would say "no". In general, we don't need quantum mechanics to explain most systemic properties.

If it is asking if we can explain consciousness by means of physical theories, the answer is "yes". This is the same situation that was discussed in section **??**.

If it is asking if we need new fundamental theories to explain consciousness the answer is "no". We have theories but what we lack are experimental and mathematical tools to properly handle it.

E. What is the relationship between the subjective and the objective?

Subjective describes means that all the information that an agent uses is always based on internal representations. This provides a varnish of self-specificity to any piece of knowledge.

Objective has no sense unless both agents are equal and posed in the same situation. Objectivity means the same information flows inside the agent in the same environmental conditions. Similar agents will perceive similar things in similar situations. This is the nature of objectivity.

When two similar agents (or identical, in the case of identical twins or twin robots) perceive similar things (have similar mental representations) in similar situations, the simpler explanation is that the world outside them is the same. We can say that they share objective information. This is however the simplest explanation of the scenario. Other alternatives exist, but this does not hamper that nature of objective information.

Information can flow from the outside to the inside of the agent, being different in origin or made different by the information representation process. The origin of subjective information can be traced back to the source of information (external or internal to the agent) or to the process of acquisition, that is always affected by agent structure (this is the Kantian epistemological view of the categories).

This phenomenon is the same for biological or robotic entities. There is no difference in perception processes except related with implementational issues.

F. To what degree are animals conscious?

G. Can machines be conscious?

Obviously. Not only they can be conscious. They are conscious. If you can look at a complex control system working in an autonomous task, you will discover

H. How does consciousness fit into the social and cultural order?

Altered consciousness states can be achieved without the use of drugs. Andrew Weil [**?**] "Hings come from within; they are simply triggered by external agents in the right conditions".

I. Are naturalistic accounts of consciousness circular, given that experience of nature is a construct of the mind?

VI. Conclusions

Research on intelligent control has reached an stagnation point because it has been mainly focused in the study of specific information technologies

We are progressing in artificial self-consciousness because we need it for achieving autonomy.