

Characterizing Human-Like Consciousness: An Integrative Approach

Ben Goertzel

OpenCog Foundation, Hong Kong, China
ben@goertzel.org

Abstract

Synthesizing concepts and findings from a number of recent models of human consciousness, a unified model of the key properties characterizing human consciousness is outlined. Six key properties are emphasized: **Dynamical representation** of the focus of consciousness, **Focusing of energetic resources** and **focusing of informational resources** on a subset of system knowledge, **Global Workspace dynamics** as outlined by Bernard Baars in his cognitive theory of consciousness, **Integrated Information** as emphasized by Tononi, and **correlation of attentional focus with self-modeling**. It is proposed that the extent, and relative importance, of these properties may vary in different states of consciousness; and that any AI system displaying closely human-like intelligence will need to manifest these properties in its consciousness as well. The “hard problem” of consciousness is sidestepped throughout, via focusing on structures and dynamics posited to serve as neural or cognitive correlates of subjective conscious experience.

Keywords: Consciousness, Information Integration, Global Workspace, Nonlinear Dynamics

1 Introduction

These days, unlike a few decades ago, consciousness is a significant topic of research in psychology, neuroscience, philosophy and other fields. However, there remains no scientific consensus on how to define or conceptualize consciousness, let alone on how to quantitatively measure it, or formally model its structure or dynamics. Among other open questions, there is no broadly accepted way to measure the degree of consciousness displayed or experienced by a system (be it a human or other animal brain, or an robot or other AI) during a certain interval of time.

I will not aim, here, to address the foundational question of “what consciousness fundamentally is.” Instead, the question I will focus on is: **What are the important properties specifically characterizing human, or human-like, consciousness?**¹

¹Just as physics has told us many interesting and useful things about the movement of objects without resolving all the core philosophical issues regarding the nature of space and time, I believe we can come to many valuable conclusions about consciousness without first needing to resolve all related philosophical perplexities.

In the longer paper [13]², the key conclusions of which this brief conference paper summarizes, a number of contemporary analyses of human consciousness are analyzed from this multifactorial perspective: Baars' Global Workspace Theory [3, 4] and the LIDA software system that partially embodies it [23]; Tononi's Integrated Information theory [26]; Goerner and Combs' analysis of consciousness in terms of nonlinear dynamics and energy minimization [10]; Tart's theory of states of consciousness [25]; and the analysis of consciousness in terms of reflective self-modeling [19, 20, 12]. It is argued that these theories, diverse on the surface, are actually elucidating different aspects of the same complex underlying human consciousness process. This brief conference paper summarizes the main conclusions of the analysis given there.

What is proposed here is to explore the possibility of measuring the degree of consciousness displayed by a human or human-like system, using multiple factors derived from these multiple theoretical perspectives. Six key factors relevant to measuring human-like consciousness are summarized as: 1) Dynamical representation of the focus of consciousness; 2) Focusing of energetic resources on a subset of system knowledge; 3) Focusing of informational resources on a subset of system knowledge; 4) Global Workspace dynamics; 5) Integrated Information, perhaps as quantified by Tononi; and, 6) Correlation of attentional focus with self-modeling.

The optimal ways to quantify all these phenomena are not yet clear. What is argued here is that, if one wishes to quantify the degree of consciousness of real-world systems, this is the right way to proceed – i.e., by identifying and then quantifying multiple aspects of the multifarious, multifactorial dynamical process that is human-like consciousness.

1.1 Hard and Possibly Less Hard Problems Regarding Consciousness

David Chalmers [6] famously distinguished the “hard problem of consciousness” from other issues regarding consciousness – where what he meant by the “hard problem” was, in essence, the problem of connecting subjective experience (the “raw feel” of consciousness, sometimes referred to using the term “qualia”) with empirically observable factors. According to my own understanding, in the current ontology of intellectual disciplines, this “hard problem” is a philosophical rather than scientific problem. My reasoning is that science, as currently understood, is focused on prediction and explanation of measurements that are observable by an arbitrary observer within a community; whereas subjective experience, by its nature, is not observable by an arbitrary observer with a community.

The topic addressed in this paper is *the cognitive and neural correlates of human and human-like consciousness* [18]. I believe this topic can be explored quite thoroughly without making any commitments regarding the “hard problem.” However, it would be dishonest of me not to clarify that I do have a personal and intellectual position on the “hard problem.” In particular, I tend to agree with Chalmers that that some sort of panpsychism is probably the right answer – i.e. I tend to view consciousness as a property that everything in existence possess to some degree and in some form. Analytic philosopher Galen Strawson [24] has argued strenuously and rigorously that any other perspective is logically ill-founded; there is no consistent, sensible way to view consciousness and the physical world as separate but interacting entities. But if one accepts this, there remain difficult questions regarding why particular physical entities are associated with particular sorts of conscious experience. The philosophical and scientific aspects of panpsychism have been explored in detail by many others [24, 7, 2] and I will not repeat those discussions here. The central ideas in this paper are not predicated on the panpsychist perspective, so I mention my orientation toward panpsychism here mainly to point out that

²See http://goertzel.org/goertzel_consciousness_review.pdf

Download English Version:

<https://daneshyari.com/en/article/486319>

Download Persian Version:

<https://daneshyari.com/article/486319>

[Daneshyari.com](https://daneshyari.com)