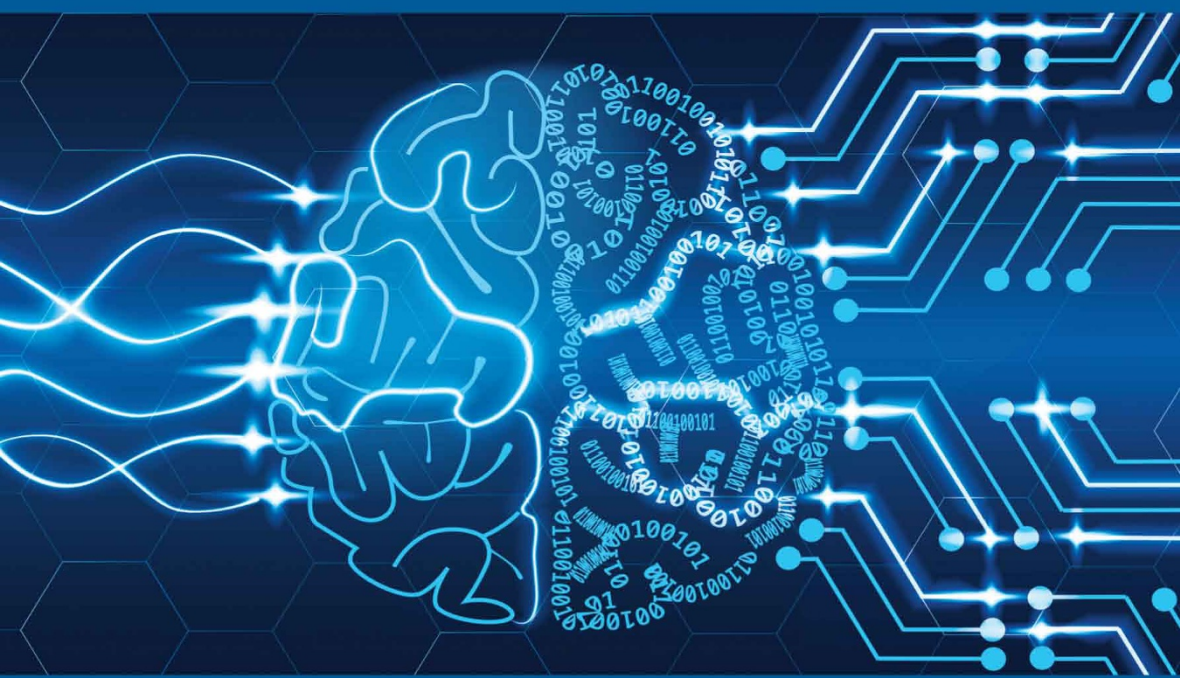
COMPUTER ENGINEERING SERIES

Beyond Artificial Intelligence

From Human Consciousness to Artificial Consciousness

Alain Cardon



ISTE

WILEY

Beyond Artificial Intelligence

Series Editor
Abdelkhalak El Hami

Beyond Artificial Intelligence

From Human Consciousness to Artificial Consciousness

Alain Cardon

ISTE

WILEY

First published 2018 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

ISTE Ltd
27–37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

Library of Congress Control Number: 2018947906

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN 978-1-78630-359-2

Contents

Table of Definitions

Introduction

Artificial intelligence is concerned with the development of computer systems that simulate human reasoning when they are applied to the domain of rational knowledge. More specific subdomains are structured by ontologies, which enable the development of systems that use this knowledge with great subtlety when questions are posed to them. This is true today of all computers and small portable devices that enable communication via the Internet on countless websites. All of these systems are therefore made to replace specialists and to help humans with their endeavors. Evolution has led to a connection between computer science and the physical, especially the electronic, which has made it possible to introduce rational behaviors into physical systems whose behavior is thereby rendered autonomous. This is how robotization has developed and continues to progress. The human being considers themselves as the pre-eminent creator, supervisor and decision-making user of these systems. This is no longer the case, since the user of a tablet or smart phone is not on their tablet or smart phone but in the device's native environment. These devices can communicate autonomously via a Hertzian network with remote systems and can make recommendations that were absolutely not requested, all while refining the user's consumer profile.

And they can do much more. These computerized systems, all of which are systems with processors and memory, can be equipped with the ability to generate forms of intentional thoughts, to have desires

and needs, and to inundate a human user in sets of procedures that they can no longer control, that are beyond them. These systems can be equipped with a psyche similar to the human psyche.

That is what this book intends to show: how the architecture of a human psychic system can be structured in an organizational approach, how a human being generates thoughts and how those thoughts then become what they feel; it then aims to show how and with what types of computer component this psyche can be transposed to transform it into a computer system that expresses an artificial consciousness. Thus, we will see how the unconscious, preconscious and artificial consciousness are structured and organized, and how all of that is brought together, with respect to information and energy, with a fourth instance: the organizational layer.

The model of the human psychic system that we will present is founded on an approach that unifies both the bottom-up and top-down approaches. The bottom-up approach considers the system to be made up of many small, highly connected parts and asks how it generates representational forms concerning the sensation of corporeality and especially the representation of symbolic evaluations of real-world objects at very high linguistic and conceptual levels. The top-down approach begins from ontologies of knowledge about everything we know how to represent cognitively and asks how to define the hierarchies of systems that express all of the categories of this knowledge from all points of departure. The unification of these two approaches is organizational and amounts to developing a system that deploys the same kind of morphologically and semantically structured components that define both foundational forms as well as those of great conceptual scope, and which ensure – especially on their own – control over multiple levels like an organizational layer.

And finally, we will see that the development of a model of the artificial psychic system by substituting the human psyche is a scientific approach that precedes building a technology for autonomous systems, and adopting a constructivist and organizational view will allow us to clarify certain characteristics of the human psyche. Science cultivates knowledge that can be shared with all

disciplines and also makes it possible to ask ethical questions about its achievements. The development and subsequent exploitation of artificial psychic systems that are equipped with intentional consciousness must necessarily raise questions concerning potential uses or even the justification of a decision not to build such systems. Therefore, the ethical question concerning the potential applications of artificial consciousnesses must now clearly be asked.

1

The Organizational Architecture of the Psychic System and the Feeling of Thinking

We are going to present an architecture of the human psychic system by adopting an organizational path that considers the psyche as a highly dynamic idea-generating system that operates continuously at different rates and structures its components on several scales to generate forms, stable for a very short time, that will be understood as the forms of thoughts. In such a framework, the question is how should such a system, made up of multiple active components on multiple scales, be designed so as to permanently produce perceptual and ideational representations of the things of the world using its abilities of naming and language abstractions? We should take into account that the system generates representations of multiple things at multiple levels and in multiple situations, allowing the human to understand reality so that they can act with a high degree of behavioral autonomy. The dynamically, spatially and temporally organized conception of representations will then be the major characteristic of the system, which is, in the end, a generator of complex constructions, usually intentionally, with highly dynamic memorizations. And we will show that the understanding of mental representations always takes place in a specific setting that brings together the psychic system's instances given certain characteristics, which we will call the mental landscape of the psychic system.

Beyond Artificial Intelligence: From Human Consciousness to Artificial Consciousness,
First Edition. Alain Cardon.

We are going to propose two models. The first model will be based on the components that carry meaning, the dynamic union of which defines the characteristics of all thoughts generated by organizing itself by means of specific elements of control. The second model will bring together the components of meaning with those of control in a unique expression that will then be morphological. The second model will represent the generative use of continuous thought-generating constructs made up of multiple aggregates of neurons connected by multiple dendrites, which produce active, emergent conformations so that the system can feel them for itself, based on a highly specific system of self-control.

1.1. The problem of the study of thought

We approach the design and generation of thoughts by taking an interest in the precise architecture of the psychic system. We are speaking about the "psychic system" and we will therefore adopt a dynamic systems modeling approach. But is it common to consider that which effectively generates thoughts as a system? In this domain, the word "system" is often troubling, because it implies – to those who are unfamiliar with dynamic models and their morphological characteristics – a reduction to mechanical and automatous features, which is obviously not acceptable in the case of the psyche. Furthermore, the position of considering the functioning of the generation of thought as that of "some type of system" is unacceptable to those who have the ability to think with immanent features engendered by an infinite source.

We refer to **idea representation** as a form of experienced thought concerning any subject. The brain continuously generates such representations by producing a series of themes of varying duration, some almost instantaneous, others whose duration depends on an intentional focus on perceived or defined subjects.

When we consider the generation of thoughts as the output of a system, we must necessarily situate the model on a certain level that cannot be reduced to the cellular level, which is the level of the minimal physical substrate. We should assert that this system is

limited with respect to its effective operational components and its potential for action and interaction, even if these limits are extremely large. We should take into consideration that this system emerges with a certain form, but develops and grows in size and organization in accordance with what is permitted by its architectural process. It continually modifies itself as it is used, and almost continuously, although at different speeds, produces idea representations with finite but multiple characteristics from its emergent states, which lead to behavioral effects ranging, for example, from movements to spoken and comprehended speech. It is an organizational system that modifies its morphology in its running, that sometimes deteriorates and that, in the end, dies with the physical host that shelters it, the human being.

This type of thought-generating system will never be a conventional state-based system, with an initial and a final state for each thought produced; this type of system would be reductive and even absurd in this case. Instead, it would be a system that is continually formed from an ensemble of active dynamic components with variable lines of potentiality and increasingly experienced emergent representations. A very organizational specific, high-level set of processes that imposes multiple constraints is required to arrange the components of the system and to transform it into an organization that will be conscious because it is experiencing the generated thoughts. A conscious event is therefore an organizational act, strictly effective for the set of components constituting the system, which puts them into a particular global state that is able to be experienced. And such an act, which does not occur by chance, must have a more or less precisely predetermined target; it has a duration, constraints, a scope and it has a global substrate at its disposal as the natural result of the operation of the system, which engenders continuous learning and development.

We assert that the generation of thoughts is the organizing process performed by brains when they are functioning, which we will refer to generally as building **experienced perceptible representations** concerning a great number of things in the world. This is what is usually referred to as “moments of experienced consciousness”. The notion of representation that we will use here is that of a complex and

completely dynamic appraisal of a constructed form, which *can be taken as* its targeted object, which is, itself, a particular thing that is understood by the system. We will refer to C.S. Peirce's triadic signs to clearly understand the meaning of the verb "to be taken as" that we are using here [PEI 84].

This kind of thought-generating system is obviously very difficult to conceive; it is completely different from a mechanism that correlates its output with its input and that operates by passing through a series of predefined states, such as in a stateful system. But it is still a system; in fact, it is a system of systems made up of multiple, strongly interconnected, dynamic processes operating at different levels that are interdependent in several ways and at several spatial and temporal scales. This system, on the fundamental physical level, activates multiple neurons via the activity of their dendrites and expresses the physical occurrence of the transmission of information flow and energy transfer. The system activates and expresses the surges of activity of processes, which we can understand in the computing sense of the term; surges in the process of neuronal actions that are complementary and especially those that occur in parallel. The very important concept of **co-activity** indicates that all actions from an emitter of information or energy modify both the receptors and the emitter itself because of this emission. This is an action that transforms the emitter and the receiver via the transfer of information or energy. The system constructs its own inputs by adapting information coming from the body's senses and endlessly constructs conscious events concerning something that was more or less intentionally targeted.

These specific configurations of the system are always ephemeral and they are produced according to the constraints that are innate or acquired because of the system's operation and the regulation of its corporeality. And these configurations will be – which is the chief attribute of the system – felt by itself, and will experience them while modifying them and memorizing them to use later to produce subsequent conscious events.

By adopting this position concerning the conception of the system, we are situated in the theory of thought generation according to a

constructivist approach by proposing an architecture that will allow for its transposition into the artificial, by situating us in the universe of swarms of constantly reorganizing processes, manipulating symbols and measurable values, constructing the organization of very dynamic structures of active elements for themselves and joining forces with each other. This is the standard position for a modeler who seeks to understand how forms as complex as ideas can be represented in the domain of verifiable knowledge and how ideas exist in and of themselves – that is to say, before they are projected into the space of words expressed in language via the production of sentences identified by sounds and symbols using grammars.

So what is a thought? What form does it have, this thing that is so real and so commonplace, so physical and yet, it seems, so hard to grasp? What is this space where it is made, initiated, expressed and memorized by altering the structure of its deployment space in order to memorize and to create others? What is it, this thing that makes it possible for the living organisms who produce them and use them to partially understand the world that surrounds them, to predict events, and, also, sometimes, to question their own existence? How can we explain the scale of that which is thought by the brains of organisms that are so evolutionarily different, and that are also characteristic of its evolution?

1.2. The interpretation of neuronal aggregates

To be able to develop a model of thought generation in a system, it is necessary to precisely clarify the characteristics of the approach as guided by its architecture. Such an approach is based on observations made in neuroscience, which analyzes neuronal activity using photographs of their energy traces, but we must also clarify what must be the architecture of the system, which is principally based on the very organized processing and manipulation of multiple sources of information. There is, in reality, a countless amount of information transmitted at the synapse level, but the understanding of the production of ideas employing words, for example, will be situated at a different level than that of the synapses. We must orient ourselves within the definition of the different architectural levels of the system

that generates and manipulates information flows, which must possess traits on the level of knowledge in order to construct dynamic forms that will become the conformations of generated and felt ideas.

The neuronal system operates on the level of production in parallel with multiple neuronal signals, which form, via their associations and aggregations, a very complex unit that can be interpreted as a structure of combined dynamic forms, a structure made up of activities and information exchanges carrying a certain level of cognitive awareness. Every organization of these dynamic forms becomes stable for a very brief moment to form a conceived thought that will be understood. The unit under consideration is therefore the production of combinations of forms of activities, and in fact of morphologies of informational and energetic forms, which combine, associate, converge and modify each other, producing a stabilized, dynamic structure for an instant, which makes it possible for the thought to thus be perceived. This is the physical generation of every produced thought, when we consider thought at its tangible level in the brain.

A thought is formed from numerous meaningful characteristics that are understood, with some characteristics being important and others secondary, contextual, associated or even opposed. The number of these characteristics is important, but it remains finite and understandable on the cognitive level. We assert that these characteristics are represented by the action of significant groups of neurons that we will call **significant neuronal aggregates**, which communicate interactively, and that these groups are interpretable as dynamic forms containing the information for generating the significant characteristics of thought. These neuronal aggregates become active with each other when asked to establish these relationships. They then activate each other at larger scales to form aggregates of aggregates, which will become the form of the expressed emergent thought. It is a question, in the model, of defining these aggregations, clarifying how and why, for what reasons, and in what qualitative contexts they can create themselves.

The consideration of what a thought is at the level of the physical substrate that permits its formation amounts to the assertion that it is an organization of complex combinations of deployed forms that

communicate, which implies that all thought is defined by the following considerations.

What is a thought?

> A thought is:
>
> – an essentially dynamic, complex physical element made up of energy and information flows from neuronal aggregates that are deployed simultaneously at several scales;
> – a dynamic construct using the memorization of the characteristics of certain forms that have already been produced;
> – a dynamic construct that expresses itself, that is used by the system that produces it so that it can be experienced, and that only lasts for the ephemeral time period as this conformation in order to perpetuate itself via additional generations of forms that will become the subsequent thoughts, in a continuous process of awakening.

Each thought is therefore a structure created in a series of produced thoughts, with strong reconstructions using forms that have been expressed and memorized. The difference in comparison to, say, a dictionary search structure is radical because there are no permanent components available, but there are reconstructions of forms that have for the most part been memorized potentially in more or less similar forms, using a memory of conformations and not a memory of components, and this occurs with the generation of each idea.

1.3. The function of the architecture of the Freudian model

We can examine five major areas concerning the study of the production of thought in the brain. There is the description of the human psychic apparatus achieved by work in psychoanalysis and psychiatry arising from the discoveries of Sigmund Freud concerning the functional architecture of the psychic system [FRE 66]. There is the work of neurobiologists, with their very fine and measured contemporary observations of neuronal activity and energy flow at the neuronal level, and even at the molecular level. There is the work in the representation of knowledge and reasoning undertaken by

cognitive science and artificial intelligence research by relying on linguistic analysis. There are recent discoveries in mathematics and computing regarding the modeling of complex systems, with theories of morphogenesis in complex systems.

And there is also, obviously, philosophy, which offers very profound reflections on what thought is, what it can mean and how it can engage in self-examination [LÉV 71]. But very few researchers have focused on the synthesis of these domains with regard to a precise topic: the organizational understanding of a system having the ability to generate the forms of thought, with the goal of developing a complete concept of the system that produces what we call the experience of perceived consciousness.

We will say that a thought is a **representation** of something precise; it is constructed, felt and assessed in order to be used and reused, and it is systematically engaged in producing other thoughts. We should specify what we mean here by representation, which is not a simple symbolic component signifying a thing, but rather the dynamic construct in the neuronal system that has created and guided a process of assessment with regard to real things that will be expressed and understood. We assert that thinking amounts to generating sequences of such felt representations with regard to the elements of reality that the system can conceive of, with attributes provided by what the architecture allows. Here is the constructivist definition that we propose for the concept of representation.

Constructivist definition of the concept of representation

> A felt **representation** is the spatial, energetic and informational generation in the neuronal system of the conformations of an organized surge of a number of elements constituting a precise structure for a brief instant. The surge will appear in the form of an internal organization of neuronal aggregates constituting a spatial conformation, formed at the minimum level by strongly connected neurons themselves constituting connected aggregates. This is a dynamic organization, evaluated sensitively by the system when it is constructed and available to be understood. This representation will be understood upon

> examination by components representing structural forces trained by multiscale actions on the physical and/or informational components of which it is constituted.
>
> We can say that it can be taken to represent things in the world because of some of its morphological characteristics, which will always be connected to the type of thing being understood. It represents, designates and expresses a real or an abstract thing, by its aspects and its characteristics, and always at a number of scales. This representation, or representational construct, is designed to transform itself into another that is more or less different, which will be the subsequent representation, thereby constituting the flow of generated representations forming thoughts, and which is impossible to interrupt. A representation is therefore not a simple functional state but a dynamic structure of informational forms that is constructed in a continuous and ever-changing procession while it is being consciously perceived.

We should clarify that the process that produces representations contains the following active and organizational components:

1) there are a large number of components of a substrate based on neurons and synapses that constitute the fundamental components whose function is to activate and connect themselves;

2) there are actions aggregating these fundamental components into spatial, energetic and informational forms having certain characteristic qualities;

3) there is an organizational action continuously aggregating all of these forms in accordance with the different spatial and energetic scales in order to build the representation. Each representation on a certain subject has a spatial and energetic conformation particular to the subject at hand, and its conformation can thereby be geometrically described by the constructive and mobile movement of conformations of physical structures having energy and conveying information;

4) there is the action of certain fundamental forces acting on the construct of that representation to extract the global meaning from it with certain important characteristics, principally when it manages to

become coherently organized, which is the conscious act that experiences its representation. The conscious experience of something is therefore a dynamic comprehension that expresses a multicharacteristic sense when the representation becomes clearly understandable, and the psychic system must be seen as an essentially perceptual and self-controllable system that is centralized for the perception of things by the human being;

5) there is a general process that leads to the use or elimination of this representation, which is activated to generate the subsequent representations in a process of continuous activity.

We should note that the process of producing the representation, which is dynamic and mobile on the energy level, influences the state of the representation itself, because it is ceaselessly modified by being constructed from relationships between its components and the targeted intention, and this property endows this process of construction with very distinctive characteristics. These characteristics are instructions for multiple actions that take place and are assessed, as always, on multiple levels. The scientific approach consists of considering this creation from geometric, dynamic and cognitive perspectives, which makes it possible to identify its characteristics in a measurable domain.

Having the idea of something specific therefore amounts to producing a felt representation indicating the aspects of the thing under consideration via an internal construction expressing the characteristics of the thing and possibly an assessment of these characteristics. This is therefore a process of construction and of the feeling of that construction, and it is certainly not a precise state that will be achieved as having been completely predictable. This representation is essentially dynamic; it is a set of activities among specific components designed to produce the representation. This set is, on the one hand, spatial, as it is always situated in several areas of the generating system, beginning with an excitation of aggregations of fundamental components, and is, on the other hand, temporal, because it is deployed and only lasts for a limited time, as it is driven to evolve or to transform by a constant, internal, continuous-layer type of control.