Ontologies as Backbone of Cognitive Systems Engineering

Ricardo Sanz[1] **and Julita Bermejo**[1] **, Juan Morago**[1] and **Carlos Hernández**[2]

Abstract. Cognitive systems are starting to be deployed as appliances across the technological landscape of modern societies. The increasing availability of high performance computing platforms has opened an opportunity for statistics-based cognitive systems that perform quite as humans in certain tasks that resisted the symbolic methods of classic artificial intelligence. Cognitive artefacts appear every day in the media, raising a wave of mild fear concerning artificial intelligence and its impact on society. These systems, performance notwithstanding, are quite brittle and their reduced dependability limips their potential for massive deployment in mission-critical applications —e.g. in autonomous driving or medical diagnosis. In this paper we explore the actual possibility of building cognitive systems using engineering-grade methods that can assure the satisfaction of strict requirements for their operation. The final conclusion will be that, besides the potential improvement provided by a rigorous engineering process, we are still in need of a solid theory —possibly the main outcome of cognitive science— that could sustain such endeavour. In this sense, we propose the use of formal ontologies as backbones of cognitive systems engineering processes and workflows.

Keywords: Cognitive systems; trust; dependability; engineering processes; autonomy; life-cycle; ontology.

1 INTRODUCTION

These days we are seeing in the media a continuous flow of reports about self-driving cars, mobile phone natural language assistants or machines that win at games traditionally considered reserved to humans (e.g. Go or Texas hold 'em poker). Artificial intelligence (AI) seems to be re-flourishing and this is raising a global awareness of its potential and a global concern of its risks. The recently created *Partnership for AI*[3] has been "Established to study and formulate best practices on AI technologies, to advance the public's understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influences on people and society."

The current flourishing of AI is characterised by an availability of high performance computing platforms that has opened an new opportunity for statistics-based cognitive systems. These systems can perform quite as humans in certain tasks that resisted approaches based on the symbolic methods of classic artificial intelligence and the formal representation approaches of cognitive science. More or less cognitive artefacts are appearing every day in the media, raising a wave of mild fear concerning artificial intelligence and its potential impact in society —both positive and negative. However, from a real-world systems usability perspective —the perspective of systems engineers— these systems are still quite brittle and their reduced dependability is seriously limiting their potential for massive deployment in mission-critical applications. A clear example are autonomous cars, where this intrinsic brittleness when deployed in real roads is limiting their widespread deployment.

In this paper we will analyse the issue of the *engineering of dependable cognitive systems*. A transition from pre-engineering to true-engineering methods is in need. The construction of a cognitive system (CS) in the past has always been a form of craftsmanship and not an instantiation of proven, repeatable engineering processes. Most CSs have not gone beyond the research phase. In a sense, all CSs that we have seen are one-of-a-kind systems. The customer-oriented repeatability of engineering methods has not been manifested at all in the domain of CS construction. Even the reuse of "commercial" cognitive technologies as IBM's Watson is still a kind of hacking. This being said, the achievement of a systematic use of engineering-grade methods for CS implementation is very important, because it could enable the strict fulfilment of user-centric requirements for their operation. This is a need for real-world deployable systems, from the dual perspective of their *capability* and their *dependability*.

This is a position paper in which we analyse the role that ontologies can play in this transition. Anticipating the final conclusion of our analysis we believe that, to achieve the potential improvement provided by using a rigorous engineering process to build CSs, we are in *need of a deeper, unified cognitive science* —a solid theory of mental processes— that could sustain such endeavour. Engineering methods based on this science of the mind will lead to the synthesis of the two classes of engineering assets that are necessary for CS engineering (CSEng): *design patterns* —structural/behavioural aspects for cognitive architectures— and *ontologies* —concepts to bind i) the minds of the engineers and system stakeholders; ii) the mind of the engineer to the CS under construction; and iii) the mind of the CS to its world and the world of its user.

The paper is structured as follows: first we analyse the concepts of *system*, *mission* and *cognition* in the domain of CSs; then we provide an overview of the engineering task of building CSs; after this, we offer a brief analysis of fundamental architectural aspects of these systems; a section on the question of dependability, resilience and trust follows; next a discussion of the need of CS-specific engineering methods; finally we include the core analysis of the need of a solid theory of cognition —based on patterns and ontologies— to support the engineering life-cycle of CSs. The paper ends with some conclusions and future perspectives.

---

[1] Universidad Politécnica de Madrid, Autonomous Systems Laboratory, Spain, email: ricardo.sanz@upm.es

[2] TU Delft Robotics Institute, Delft University of Technology, The Netherlands c.h.corbato@tudelft.nl

[3] https://www.partnershiponai.org/

2 ENGINEERING COGNITIVE SYSTEMS

The issue that we stress in this paper is the need of having systematic engineering methods to build custom cognitive systems and the critical role that ontologies will play in achieving this.

2.1 Systems and Missions

AI is artificial not because it is not natural, but because, as described by Simon [63], things are *built to satisfy a purpose*. Real-word systems[4] should be useful to people; their users. Users expect that the systems will be able to fulfil their missions as planned.

When we ask Siri to find a nearby Spanish restaurant we expect i) to get an answer; ii) that the restaurant suggested is close to our position; and iii) that it is indeed a Spanish restaurant and not a Mexican one or a New-yorkean tapas bar. Good engineering is always concerned about building *systems* that reliably fulfil their *missions* [39].

The use of AI techniques in any kind of systems help these systems deal with complex problems and situations. We call these systems *cognitive* because they "know", *i.e.* hey exploit some form of knowledge[5] in the performance of their missions[6]. AI systems have been able to progress in problem solving steadily [46], overcoming many of the predicted limits [24]. Most of the envisioned limits of AI are related to aspects of the human mind that are considered particularly peculiar like creativity and consciousness. These and other, simpler aspects of human minds are the essential focus of Cognitive Science (CogSci) and are explored in the heterogeneous ways and domains that characterise CogSci [71].

Computer implementations of CSs have always been focused on essentially two parallel but different purposes: i) the evaluation by computer simulation of cognitive theories of the biological mind (esp. human); and ii) the construction of artefacts able to intelligently deal with their complex worlds. In some cases these two threads are mixed —e.g. when implementing cognitive anthropoids— leading to confusion about the purposes and the degree of success of the development of the CS.

Precise statement of system goals and operational requirements is a strong principle (almost a dogma) of systems engineering and is badly needed in the context of CS construction; otherwise projects get lost into disparaged explorations sometimes concerned about the mission of the system, sometimes about properly mimicking human features, and sometimes about the exploration of the elusive landscapes of cognitive capabilities.

In the domain of Cognitive Systems Engineering (CSEng)[7] it is necessary to focus on the mission that the CS shall perform, trying to avoid vagaries into its human resemblances. On the other hand, mimicking humans is a perfectly valid class of mission. So, we can indeed undertake the mission of modelling the human mind using engineering-grade methods, but this requieres a precise statement of this goal and a method of verification, *i.e.* his needs *an objective test of human-likeness*. However, this specification of an *objective* Turing test is however still a dream.

In essence, CSEng seeks to properly use artificial intelligence at the service of improving mission-level capability. Cognitive systems are able to see, perceive objects and affordances, making good decisions and acting properly in the world in the service of a pre-specified mission —all this driven by knowledge available to the agent or gained through the senses. Besides having intrinsic capability —e.g. the capability of following a path— cognitive systems must also be dependable. Trust and usability rely both on dependability.

Robustness —the capability of tolerating disruption— and resilience —the capability of functional recovery— [35] are hence critical aspects for real-world systems and one the main negative aspects of research-grade CSs.

2.2 Engineering Life-cycles for Cognitive Systems

The expression *Cognitive Systems Engineering (CSE)*[8] has been used by other authors to refer to the engineering of systems that include humans performing cognitive tasks. In essence it is used to refer to a special discipline of *human factors* in systems development that addresses the design of human-cognitive socio-technical systems (a system in which humans are the providers of essential capability related to perceiving, evaluating, deciding, planning and executing [36, 55, 72]. CSE is in essence the system analysis and design effort necessary to support the cognitive requirements of human work inside socio-technical systems [41, 48].

In this paper, however, we will use the term *Cognitive Systems Engineering* to refer to the wider discipline of building systems that are cognitive (e.g. including humans, AIs or both) in support of the *mission itself*. Inside this large domain, we are specifically interested in artificial systems that are cognitive by themselves, not by being in relation with humans who are the thinkers. Obviously, the characterisation of what is cognitive and what is not deserves special attention [1], and, while it is not the central aspect of this paper, it will be necessarily addressed later in the section on the need of a deeper cognitive science (see Section 3).

Cognition is a general capability (*i.e.* ot exclusively human or biological) that can be provided by machines, hence the inclusion of humans is not a necessary condition for a system to be cognitive. Nor it is necessary for a system to behave like a human to be cognitive. In our work at the UPM Autonomous Systems Laboratory, we are specifically interested in CSs that do not depend on cognitive capabilities of other agents —esp. humans— to fulfil their missions. We are interested in autonomous CSs. Cognition in artefacts can span a wide range: from minimal cognition [8] to human-like performance in specific tasks [66].

This wide spectrum of cognition apparently implies that there is no possibility of having a single, established engineering process and cognitive system life-cycle. This process and life-cycle would simplify the construction process and provide improvements concerning the predictable fulfilment of user requirements. A systematic engineering methodology for CSEng is needed [14] and it shall be based on the right architectural patterns and reusable assets [19].

Figure 1 summarily depicts a part of the engineering processes and system assets that are performed and appear along the life-cycle of a cognitive system. It shows the two/three major phases in a CS life-cycle: i) engineering —design and construction— and ii) operation.

---

[4] We use the term *real-world* to refer to systems that are deployed to serve a specific purpose, in contrast to research-oriented systems.

[5] While knowledge and belief are obviously different things, for practical reasons in this paper we do not make a strong distinction between them.

[6] Note that this conception of *cognitive* somehow departs for a widespread understanding of the term that equates it with *having some resemblance of human mental processes*. This conception of cognition is excessively anthropomorphic and lessens the possibility of advancing in a general cognitive science.

[7] The term *Cognitive Systems Engineering* has been used by other authors [36, 55, 72] to refer to systems that include humans doing cognitive work —see next Section— but we use this term here to refer to the engineering of (maybe artificial) cognitive systems like those mentioned before —e.g. Siri or the Tesla autonomous car.

[8] Note that this is different from CSEng described before.

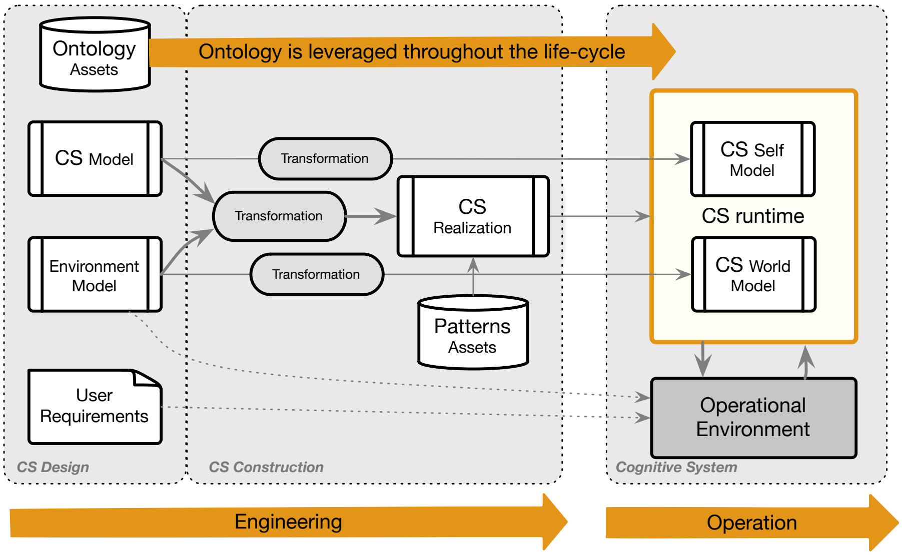

Figure 1. *The cognitive system engineering life-cycle. Systems concepts support the whole process, providing a framework for analysis, design, construction and operation of targeted cognitive systems. Ontologies capture the concepts in the mind of engineers and their formalisations can be used to synthesise parts of the architecture of the target cognitive system including the representations used for its mental content: perceptions and reflections of the CS's world.*

Resilient CSs use runtime models of their operational environment and models of themselves to sustain their activity. The CS is built from assets realised by transformation of the CS engineering models and reusable assets in the form of ontologies and patterns. In this paper we focus on the role of ontologies. Details will be given later (in Section 4).

2.3 Cognitive Systems Designs

The literature on cognitive systems is full of examples of architectures for cognition [43]. Some of them are focused on very specific tasks while others are postulated as almost universal [42, 4]. Some of these architectures come with associated engineering methods that can be applied to build custom systems using the reusable assets that the architecture provides [2]. However, from a perspective of general CSEng, these architectures come with excessive a priori commitments that limit the decisions that a cognitive system architect may take [62].

There is a need of lowering the sophistication and completeness of descriptions of generic cognitive systems, going back from full creatures [12, 22] to focused functional patterns [3, 15, 49] that can provide specific capabilities (e.g. recognition, learning, metacognition, etc.). From the CSEng perspective, there is a need of moving back from the search of *unified theories of cognition* proposed by Newell [50] into a search of *unifiable patterns of cognition*.

This means that cognitive systems designs shall be described at two levels: i) at the *whole system level* in the form of reference architectures and ii) at the *functional subsystem level* in the form of specific cognitive patterns [17, 18].

Reference architectures and cognitive patterns are the necessary design assets for the realisation of the engineering process suggested in the previous section. Note that the patterns are what provide the specific capabilities while the architectures provide the framework for their integration [60]. Note also that autonomous cognitive systems require both direct capabilities —e.g. face recognition— and meta capabilities —e.g. learning— to provide the required levels of autonomy [16]. Note also that patterns for general resilience — necessary to build up trust— are mostly based on reflection and cognitive self-penetrability of the agent itself [54].

3 A DEEPER COGNITIVE SCIENCE?

The previous analysis implies that there is a need of clarifying "cognition" from the perspective of elementary capabilities. The *unifiable patterns of cognition* mentioned earlier should be general systems' generalisations of the micro-theories that psychology has been building about cognitive phenomena in humans [51]. Stork [67] claims for a scientific foundation for engineering cognitive systems but, for the reasons explained before, this scientific foundation shall also encompass humans —and animals. We need a deeper cognitive science that can ground the rigorous engineering of cognitive systems.

Identification of common concepts between humans and machines is critical to provide a grounded integration of humans in cognitive socio-technical systems [53]. The conventional approach of providing conceptual mappings through the human-machine interface (HMI) between extremely different human/machine ontologies does not provide the constructibility, flexibility, and dynamic allocation of tasks that is needed for mixed autonomy resilient systems [23].

The real possibility of a scientific, unified cognitive science is still a theme of debate [70, 25]. There is a common thought that "cognitive science" was a different discipline to different groups of scholars, and that even fundamental aspects would remain open forever. For example, Peterson [52] bewares of unification risks, pointing at the possibility of empty relabelling of mental phenomena done by "representational cognitivism" and the practical neglect of different viewpoints (like enactivism or embodiment). In the same vein, Husbands says that "More than a decade ago, in his fine exploration of the then burgeoning cognitive science movement [30], Gardner asked if neuroscience would devour the entire field. His answer was no, for the same reasons that he could see no prospects for a single unified discipline: there will always be separate subject areas, we will always need diverse multi-level descriptions and explanations." Dale

[20] warns us that "cognitive scientists should take seriously the possibility that a single, unified framework for all of cognition is an unrealistic expectation for its diverse interdisciplinary goals and subject matter".

However, there is also a generalised feeling of an intrinsic, underlying reality of cognition that shall be captured by a deep cognitive science [59]. For example, Husbands [37] says that "solutions to the problem of how brains produce adaptive behaviour would necessarily entail (at least implicit) specifications for building artificial brains [...] being potentially useful practical devices". This points to the cognitive patterns mentioned in the previous section.

Cognitive science shall be a science of cognition and not only a science of human cognition. Obviously, human minds are specially capable in some contexts and hence their preeminence in cognitive science. However non-human minds[9] are also very relevant for a universal science, non anthropocentric view [33, 40]. This is of special importance on the coming years where cognitive artefacts will progressively take roles reserved to humans. A solid theory of mind —esp. the artificial mind— is needed to fulfil requirements and to avoid emergent, disruptive phenomena [11, 44].

The possibility of a unified science of mind may seem remote, but note that the "computer" mind model seems distant to human minds due to its digital nature. The sentence "thinking is computation" may hence be seen as plainly false. It is necessary to stress the fact that "computation" is not necessarily "digital computation". A paradigmatic example of this case are analog computers. Note also that the brain is both discrete and continuous. In the words of Schadé and Smith [61]: "Thus one single neuron possesses the qualities of a hybrid computer."

There is a strong need of finding the right framework where to build the *definite theory of mind*[10]. The theoretical frameworks we are using today to formalise system structure and behaviour are quite disjoint. For example, when reasoning about temporal dynamics of system state, consider the difference between representing change by means of temporal logic [27] or systems dynamics [45]. The models that we use are sometimes seen as mere rough approximations [64].

While it may seem difficult to achieve a theoretical agreement on a *universal theory of cognition*, most present theories are not too far from each other. The cognitive science debates are grounded not on theoretical issues but on attachments to disciplinary practices and languages, excessive targeting of full-fledged human cognition or excessive narrowing on specific cognitive tasks and feats. For example, when analysed in terms of general systems theory, classical cognitivism, embodied cognitive science or enactivist cognition are not so far from each other [47]. In fact, if we eliminate representation, what remains is mindless body; it may be performant, but only in a singular task[11]. If there is no information, there is no adaptation. We must be able to separate *what minds are* from *what happens when minds are attached to a particular environment*.

4 ONTOLOGIES AS BACKBONE

The essential question is: What are the bricks to build the edifice of the deep cognitive science described in the previous section? What is the theoretical substrate that grounds the description of architectures and patterns needed for cognitive systems engineering? The answer is simple: concepts [5, 29]. Fundamental concepts about mental phenomena will provide the necessary assets to build the cognitive capability theories and the patterns and architectures that reify them. Ontologies —as conceptual structures— will be the backbone that will sustain the cognitive systems engineering processes of the future, because they will both support the human engineering activities and the operation of the cognitive systems themselves.

It must be noted that these ontologies shall not be restricted to mind internals, but must also capture the realities outside cognitive agents [**?**]. The CS *environment* shall be formally addressed in ontological work to enable both the engineering of single-CS activity and multi-CS cooperation. Note also that these ontologies and their use must be flexible enough as to be able to accommodate differences and inconsistencies between different CS (or even between the the ontologies themselves). Ontologies shall move to use less rigid logical frameworks (capable of addressing *belief*, *uncertainty* and emphinconsistency).

Within computer science and software engineering, an ontology defines a set of ontological elements as representational primitives that can be used to model a domain of knowledge [31], [32]. The underlying idea is that ontologies provide a common vocabulary with explicit semantics [65].

Ontologies would act as a foundation for system science, where they facilitate organising the domain knowledge and formalising the different engineering artefacts or entities (e.g stakeholder, requirement, function, scenario, etc.) and their relationships, to resolve ambiguities and check for consistency and completeness [10], [38]. As examples of this approach, ontologies have been used to: assist the systems engineering process when it comes to establishing requirements [28]; drive the specification of the structural, functional, behavioural knowledge for a domain–specific system design [6]; or establish a set of common and shared concepts identified and agreed upon for autonomous systems representation and engineering [7].

When ontologies are reified as mental assets for autonomous CSs they will play both roles of defining the meaning of terms and grounding the understanding of the world (*i.e.* he CS itself and its environment including other CSs). Ontologies would act as a representation-based mechanism expressible in a computational language, to describe the different entities participating in the design and operation of the system. Concepts will then become computational elements in knowledge models that would exist through their representations in software, such as UML diagrams [26] or ontology-related languages such as OWL [68]. Concepts in engineering minds and engineering models will also be reified as physical elements in the CS or in its environment. This accounts for the modelling relation that underlies life and cognition [56]. As a consequence, the knowledge models will no longer be characterised by the usual high level of arbitrariness or difficult reuse, as it happens when knowledge is formalised as vocabularies expressed in natural language —the usual practice in cognitive science. How the meanings of concepts defined in the ontologies are understood by different components of the cognitive system is as much the focus as how the different actors in the CS life-cycle can use a common conceptualisation of the problem under scrutiny, to come up with a possible solution.

The complementary element of this ontological vision is the use of design patterns for description and engineering if system capabilities, as they provide solutions to design problems that happen repeatedly [34]. Such design patterns will act as the cornerstones of reusing CS architectures [60]. These patterns could be either archi-

---

[9] In this context, "mind" shall not be understood as a full-fledged human psyche. A mind is an information-driven controller for a system. These appear in all the biological and artificial spectrum in different degrees of capability, consciousness and self-awareness [57].

[10] A single account for human, animal and artificial minds.

[11] Arguments about the possibility of non-representational minds (e.g. [13]) are easily deflated by simple pointing at the representations used in the proposed "representationless" architectures.

tectural or domain patterns [15]. *Architectural patterns* express the structural organisation of a cognitive system, *i.e.* hey realise its architecture. *Domain patterns* describe a mechanism to solve a concrete but recurring problem in a particular context. These point towards the cognitive micro-theories that populate the literature. Note that architectural and domain patterns are not usable in isolation; in a concrete CS both shall be merged to *put mind —architecture– to a task —domain.*

It is worth pointing out that these patterns will not be used independently but having a domain ontology acting as backbone (see Figure 1). The architectural patterns will describe the cognitive system internal organisation and dynamics, based on the interactions between the ontological elements that describe the system elements themselves. Domain patterns in turn will describe the interactions among the cognitive system components, and with the environment, by using the underlying conceptualisation specified by the ontology, that represent design solutions so that the behaviour of the cognitive system fulfils the engineering requirements. Thus all system patterns will not only be specified from the ontological concepts, but eventually will become part of the ontology itself, modelling the relations and interactions between them as designed by the engineers [9].

5 CONCLUSIONS

Dale, Dietrich, and Chemero [21] defend the idea that " the 'framework debate' in cognitive science is unresolvable. The idea that one theory or framework can singly account for the vast complexity and variety of cognitive processes seems unlikely if not impossible.".

In the context of cognitive systems engineering —with and without humans— this idea is untenable for two reasons, one theoretical, one practical:

- Science advances by unification [69]. This can be considered a theoretical dogma or just a leitmotif of science. In any case, it is inalienable in the opinion of these authors because it is needed to support engineering-grade CS construction (see Section 2).
- Engineering works by using commensurable designs. There are not —there cannot be— different theoretical frameworks underlying systems of "vast complexity and variety" no matter what level of complexity they have. This does not mean that there cannot be different realms, but all them shall integrate seamlessly under a common framework. For example, mechanical engineering and electrical engineering address different domains of knowledge and practice. However, both fall under the physical framework. This is what enables the construction of solid electromechanical systems (e.g. the concepts of energy, force and torque are shared in both realms; this is what enables electrical motors moving machines).

Science and engineering are necessarily bound together under a common framework. The differences that manifest in different systems or theories shall be sought for not in the theoretical framework but in the concrete instantiations of that framework on specific entities. For example, electrical systems theory is a *single theory* even when the electrical systems in the US and in Europe have very important differences. Having a common theoretical framework implies that differences can be overcome; electrical systems can be interconnected and interoperate as a system-of-systems able to fulfil a single mission. The same can be said about cognitive systems: differences shall be found in concrete realisations, not in the theoretical concepts that underlie them.

Intense work shall be done to identify, clarify and formalise the concepts and patterns that underlie all classes of cognitive behaviour beyond the anthropomorphic trap [58]. Their expression in the form of formal ontologies will then constitute core assets for the engineering of the trustable autonomous cognitive systems of the future. The fundamental concepts used by engineers to think about the cognitive system under construction and the fundamental concepts used by the cognitive system itself in its interaction with its world —including itself— will be the same and captured in a formal representation. This will open a new world of capabilities for cognitive systems that will be able to deeply adapt to changing worlds in pursue of their missions.

ACKNOWLEDGEMENTS

We acknowledge the support of the Spanish Government through grant RTC-2016-5191-8 Sistema Autónomo de VIgilancia y SEguridad basado en multirotores (ADVISE).

REFERENCES

[1] Frederick Adams and Rebecca Garrison, 'The mark of the cognitive', *Minds and Machines*, **23**, 339–352, (2013).
[2] James S. Albus and Anthony J. Barbera, 'RCS: A cognitive architecture for intelligent multi-agent systems', *Annual Reviews in Control*, **29**(1), 87–99, (2005).
[3] Christopher Alexander, Sara Ishikawa, and Murray Silverstein, *A Pattern Language: Towns, Buildings, Construction*, Oxford University Press, 1977.
[4] John L. Anderson, 'Autonomous systems intelligence', in *Proceedings of the 1983 Annual Conference on Computers : Extending the Human Resource*, ACM '83, pp. 229–233, New York, NY, USA, (1983). ACM.
[5] A. T. Bahill and B. Gissing, 'Re-evaluating systems engineering concepts using systems thinking', *IEEE Transaction on Systems, Man and Cybernetics, Part C: Applications and Reviews*, **28**(4), 516–527, (1998).
[6] Shreya Banerjee and Anirban Sarkar, 'Ontology–driven approach towards domain-specific system design', *International Journal of Metadata, Semantics and Ontologies*, **11**(1), 39–60, (2016).
[7] Behzad Bayat, Julita Bermejo-Alonso, Joel Carbonera, Tullio Facchinetti, Sandro Fiorini, Paulo Goncalves, Vitor A. M. Jorge, Maki Habib, Alaa Khamis, Kamilo Melo, Bao Nguyen, Joanna Isabelle Olszewska, Liam Paull, Edson Prestes, Veera Ragavan, Sajad Saeedi, Ricardo Sanz, Mae Seto, Bruce Spencer, Amirkhosro Vosughi, and Howard Li, 'Requirements for building an ontology for autonomous robots', *Industrial Robot: An International Journal*, **43**(5), 469–480, (2016/09/14 2016).
[8] Randall D. Beer and Paul L. Williams, 'Information processing and dynamics in minimally cognitive agents', *Cognitive Science*, **39**(1), 1–38, (2015).
[9] Julia Bermejo, Carlos Hernández, and Ricardo Sanz, 'Model-based engineering of autonomous systems using ontologies and metamodels', in *IEEE International Symposium on Systems Engineering 2016 (IEEE ISSE 2016)*, Edinburgh, Scotland, (2016).
[10] Ralph Bogusch, 'Ontology-based systems engineering - the smart way of realizing complex systems'. Keynote Lecture, 13 November 2015.
[11] Nick Bostrom, *Superintelligence: paths, dangers, strategies*, Oxford University Press, 2014.
[12] Rodney A. Brooks, 'The whole iguana', in *Robotics science*, ed., Michael Brady, 432–458, MIT Press, Cambridge, MA, USA, (1989).
[13] Rodney A. Brooks, 'Intelligence without representation', *Artificial Intelligence*, **47**(1-3), 139–159, (1991).
[14] Dennis M. Buede, *The Engineering Design of Systems*, Wiley, 2009.
[15] Frank Buschmann, Regine Meunier, Hans Rohnert, Peter Sommerlad, and Michael Stal, *Pattern Oriented Software Architecture. A System of Patterns*, John Wiley & Sons, Chichester, UK, 1996.
[16] Nicholas L. Cassimatis, Paul Bello, and Pat Langley, 'Ability, breadth, and parsimony in computational models of higher-order cognition', *Cognitive Science*, **32**, 1304–1322, (2008).
[17] Robert Cloutier, Gerrit Muller, Dinesh Verma, Roshanak Nilchiani, Eirik Hole, and Mary Bone, 'The concept of reference architectures', *Systems Engineering*, **13**(1), 14–27, (2010).

[18] Robert J. Cloutier and Dinesh Verma, 'Applying the concept of patterns to systems architecture', *Systems Engineering*, **10**(2), 138–154, (2007).

[19] Edward Crawley, Bruce Cameron, and Daniel Selva, *Systems architecture: Strategy and product development for complex systems*, Prentice Hall,, Hoboken, NJ, USA, 2015.

[20] Rick Dale, 'The possibility of a pluralist cognitive science', *Journal of Experimental & Theoretical Artificial Intelligence*, **20**(3), 155–179, (2013/04/30 2008).

[21] Rick Dale, Eric Dietrich, and Anthony Chemero, 'Explanatory pluralism in cognitive science', *Cognitive Science*, **33**(5), 739–742, (2009).

[22] Daniel Dennett, 'Why not the whole iguana?', *Behavioral and Brain Sciences*, **1**, 103–104, (1978).

[23] A. Dominguez-García, G. Hanuschak, S. Hall, and E. Crawley, 'A comparison of GN&C architectural approaches for robotic and human-rated spacecraft', in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, pp. 20–23, (2007).

[24] Hubert L. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason*, The MIT Press, 2 edn., 1992.

[25] Barbara Von Eckardt, *What is Cognitive Science?*, Bradford Books, MIT Press, 1995.

[26] James Rumbaugh et al., *The Unified Modeling Language Reference Manual*, Addison-Wesley, 1999.

[27] *Handbook of Temporal Reasoning in Artificial Intelligence*, eds., M. Fisher, D. Gabay, and L. Vila, Elsevier, 2005.

[28] Anabel Fraga, Juan Llorens, Luis Alonso, and José M. Fuentes, 'Ontology-assisted systems engineering process with focus in the requirements engineering process', in *Complex Systems Design & Management: Proceedings of the Fifth International Conference on Complex Systems Design & Management CSD&M 2014*, ed., Frédéric et al. Boulanger, 149–161, Springer International Publishing, (2015).

[29] Peter Gardenfors, *Conceptual Spaces. The Geometry of Thought*, MIT Press, Cambridge, MA, 2000.

[30] Howard Gardner, *The Mind's New Science. A History of the Cognitive Revolution*, Basic Books, 1985.

[31] T. Gruber, 'Ontology', in *Encyclopedia of Database Systems*, eds., L. Liu and M. Tamer Ozsu, 1963–1965, Springer US, (2009).

[32] Nicola Guarino, Daniel Oberle, and Steffen Staab, 'What is an ontology?', in *Handbook on Ontologies*, 1–17, Springer Berlin Heidelberg, (2009).

[33] Richard Harper, Alex Taylor, and Micheal Molloy, 'Intelligent artefacts at home in the 21st century', in *Material Agency. Towards a Non-Anthropocentric Approach*, eds., Carl Knappett and Lambros Malafouri, 97–120, Springer, (2008).

[34] Carlos Hernández, Julita Bermejo-Alonso, Ignacio López, and Ricardo Sanz, 'Three patterns for autonomous robot control architecting', in *The Fifth International Conference on Pervasive Patterns and Applications - PATTERNS 2013*, pp. 44–51, Valencia, (May 27 - June 1 2013).

[35] Kjell Jørgen Hole, *Anti-fragile ICT Systems*, SpringerOpen, 2016.

[36] Erik Hollnagel and David D. Woods, *Joint cognitive systems: foundations of cognitive systems engineering*, Taylor & Francis, Boca Raton, FL, 2005.

[37] Phil Husbands, 'Cognitive science: Real or imaginary?: Review of the MIT encyclopedia of the cognitive sciences edited by robert a. wilson and frank c. keil', *Artificial Intelligence*, **130**(2), 191–194, (8 2001).

[38] INCOSE, *Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities, version 4.0*, John Wiley and Sons, Inc, Hoboken, NJ, USA, 2015.

[39] Scott Jackson, *Architecting Resilient Systems: Accident Avoidance and Survival and Recovery from Disruptions*, Wiley, Hoboken, New Jersey, USA, 2010.

[40] Owain Jones and Paul Cloke, 'Non-human agencies: Trees in place and time', in *Material Agency. Towards a Non-Anthropocentric Approach*, eds., Carl Knappett and Lambros Malafouri, 79–96, Springer, (2008).

[41] G. Klein, S. Deal, and S. Wiggins, 'Cognitive systems engineering: The hype and the hope', *Computer*, **41**(3), 95–97, (March 2008).

[42] John E. Laird, Allen Newell, and Paul S. Rosenbloom, 'Soar: an architecture for general intelligence', *Artificial Intelligence*, **33**(1), 1–64, (1987).

[43] Pat Langley, 'Cognitive architectures and general intelligent systems', *AI Magazine*, **27**(2), 33–44, (2006).

[44] Nancy Leveson, *Engineering a safer world: systems thinking applied to safety*, Engineering systems, The MIT Press, Cambridge, Mass., 2012.

[45] Lennart Ljung, *System Identification: Theory for the User*, Prentice Hall PTR, 2nd edn., December 1998.

[46] Stephen Lucci and Danny Kopec, *ArtificiAl intelligence in the 21st century: A Living Introduction*, Mercury Learning and Information, 2nd edition edn., 2016.

[47] Mohan Matthen, 'Debunking enactivism. debunking enactivism: a critical notice of hutto and myin's radicalizing enactivism', *Canadian Journal of Philosophy*, **44**(1), (2014).

[48] Laura G. Militello, Cynthia O. Dominguez, Gavan Lintern, and Gary Klein, 'The role of cognitive systems engineering in the systems engineering design process', *Systems Engineering*, **13**(3), 261–273, (2010).

[49] Michael S. P. Miller, 'Building minds with patterns'. ISBN-10: 0-692-54140-1, 2016.

[50] Allen Newell, *Unified Theories of Cognition*, Harvard University Press, 1990.

[51] Allen Newell and Herbert A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, 1972.

[52] Donald M. Peterson, 'The mind's new labels?: Review of R.A. Wilson and F.C. Keil (eds.), the MIT Encyclopedia of the Cognitive Sciences', *Artificial Intelligence*, **130**(2), 213–216, (8 2001).

[53] *Human-System Integration in the System Development Process: A New Look*, eds., Richard W. Pew and Anne S. Mavor, The National Academies Press, 2007.

[54] Zenon Pylyshyn, 'Is vision continuous with cognition? the case for cognitive impenetrability of visual perception', *Behavioral and Brain Sciences*, **22**, 341–364, (1999).

[55] Jens Rasmussen, Annelise Mark Pejtersen, and L.P. Goldstein, *Cognitive Systems Engineering*, Systems Engineering, Wiley-Interscience, 1994.

[56] Robert Rosen, *Anticipatory Systems. Philosophical, Mathematical, and Methodological Foundations*, volume 1 of *IFSR International Series on Systems Science and Engineering*, Springer, 2nd edn., 2012.

[57] Ricardo Sanz, 'Machines among us: Minds and the engineering of control systems', *APA Newsletters - Newsletter on Philosophy and Computers*, **10**(1), 12–17, (2010).

[58] Ricardo Sanz, 'Escaping the human constraint. towards a general theory of consciousness', in *Proceedings of the 16th Conference of the Association for the Scientific Study of Consciousness*, Brighton, UK, (July 14-17 2012).

[59] Ricardo Sanz, 'Reducing behaviorism and cognitivism', in *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, eds., M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth, Berlin, (July 31-August 3 2013).

[60] Ricardo Sanz and Janusz Zalewski, 'Pattern-based control systems engineering', *IEEE Control Systems Magazine*, **23**(3), 43–60, (June 2003).

[61] J.P. Schadé and J. Smith, 'Computers of the brain and brainmade computers', in *Computers and Brains*, eds., J.P. Schadé and J. Smith, volume 33 of *Progress in Brain Research*, Elsevier, (1970).

[62] Daniel Selva, Bruce Cameron, and Ed Crawley, 'Patterns in system architecture decisions', *Systems Engineering*, n/a–n/a, (2016).

[63] Herbert A. Simon, *The Sciences of the Artificial*, MIT Press, Cambridge,USA, 3rd edn., 1996.

[64] Herbert A. Simon and Allen Newell, 'Models : their uses and limitations', in *The state of the social sciences*, ed., Leonard D. White, 66–83, University of Chicago Press, Chicago, (1956).

[65] Ben Smith, Valentina Tamma, and Michael Wooldridge, 'An ontology for coordination', *Applied Artificial Intelligence*, **25**, 235–265, (2011).

[66] James J. Staszewski, 'Models of human expertise as blueprints for cognitive engineering: Applications to landmine detection', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **48**(3), 458–462, (2017/01/26 2004).

[67] Hans-Georg Stork, 'Towards a scientific foundation for engineering cognitive systems –a european research agenda, its rationale and perspectives', *Biologically Inspired Cognitive Architectures*, **1**(0), 82–91, (7 2012).

[68] W3C. OWL - web ontology language - semantic web standards, 2012.

[69] Edward O. Wilson, *Consilience. The Unity of Knowledge*, Alfred A. Knopf, New York, USA, 1998.

[70] R. A Wilson, 'The cognitive sciences: A comment on 6 reviews of the mit encyclopedia of the cognitive sciences', *Artificial Intelligence*, **130**(2), 223–229, (8 2001).

[71] *The MIT Encyclopedia of Cognitive Sciences*, eds., Robert A. Wilson and Frank C. Keil, MIT Press, Cambridge, MA, 1999.

[72] David D. Woods, 'Cognitive technologies: The design of joint human-machine cognitive systems', *AI Magazine*, **6**(4), 86–92, (1985).