Cognitivist and Emergent Cognition - An Alternative Perspective

Michael James Gratton

School of Computer Science and Engineering,
The University of New South Wales, Australia
mikeg@cse.unsw.edu.au

Abstract. A new perspective on classifying cognitive systems is presented in which a distinction is made based on how the world is represented, rather than the typical distinction between cognitivist and emergent approaches. It is argued that the typical classification in essence distinguishes between systems by their implementation, rather than by their properties. The alternative presented here instead focuses on how the system represents the world (if at all) and whether these representations are intelligible to the designer or the system itself. From this novel angle, existing systems are better classified and importantly a gap in existing cognitive systems research becomes evident. An outline of a well-founded cognitive system that fills this space is put forward, one which cognitive robotics is ideally situated to explore.

1 The Embodiment Divide

Cognitive robotics research aims to produce robotics systems that approach the level of cognitive ability that we as humans display, exhibiting capabilities such as perception, learning, predicting and taking action, in an autonomous manner. In *cognitivist* systems such as Shakey [1], cognition is a modular computational process underwritten by the *physical symbol system hypothesis* [2]. Alternatives such as the *subsumption architecture* [3], and those based on neural networks, dynamical systems and others were developed later as a response to perceived shortcomings of the classical cognitivist approach. These have been collectively termed *emergent systems* in surveys of the state of the art in cognitive systems (e.g.: [4]). Cognitive systems are now commonly classified along those lines: cognitivist versus everything else. While this distinction may be useful in comparing the implementation approaches of cognitive systems, it does not provide any further depth of insight into the differences between them.

This problem manifests itself as ambiguities in classifying systems by characteristics which should properly provide clear lines of distinction. For example, in the survey above, a system's "representational framework" is one such characteristic, with cognitivist systems being described as having "patterns of symbol tokens" as a representational framework while emergent systems instead have a "global system state". It can be argued however that a collection of patterns of symbol tokens taken together also constitutes a kind of global state, thus this

K.-U. Kühnberger, S. Rudolph, and P. Wang (Eds.): AGI 2013, LNAI 7999, pp. 192–195, 2013.

characteristic fails to delineate a property that successfully distinguishes between the two families.

Another characteristic, one that is commonly used to differentiate between cognitivist and emergent systems, is the notion of *embodiment* — the situation of a system in its environment such that it can not only be influenced by it via perception but can also influence it back by taking action. Part of the importance attached to embodiment stems from its perceived necessity for autonomous systems such as those sought for cognitive robotics: It is precisely the co-interaction of system with its environment that gives rise to perceiving, acting, learning and other processes of cognition and hence without embodiment these processes cannot occur [5]. Also of note is the necessity of embodiment for emergent systems, which for the most part having ad-hoc representations (if any at all [6]) are driven directly by perception. This is in contrast to cognitivist systems which, by virtue of having internal representations composed of abstract symbols, are freed from the necessity of embodiment. For reasons such as these, embodiment is a property commonly ascribed to emergent systems. However again it is possible to argue that a cognitivist system could also require embodiment — any cognitive robotics system that treats percepts as symbols is an example, and so embodiment also fails to be an effective distinguishing characteristic.

Since even a key characteristic such as embodiment fails to clearly delineate between cognitivist and emergent systems, something is clearly amiss. The traditional classification is in effect distinguishing between systems based purely on their implementation details, rather than the properties of the systems. The obvious question then arises: Which properties are useful in discriminating between different kinds of cognitive systems, in some more illuminating way other than by just the details of their implementation?

An answer can by found by noting that the notion of embodiment as a requirement for a cognitive system is in fact somewhat misdirected. The question of its necessity is actually one of whether a system has representations independent of its environment or not, rather than if the system is physically instantiated. This insight, along with well-founded arguments about the nature of mental representations from philosophy of mind and cognitive science allows at least two distinctions to be teased out. The first is the content of a system's representations. Cognitivist systems can be said to have *systematic representations*, being representations that are structured and are able to be combined in an infinite number of ways. By contrast, emergent systems have unstructured or ad-hoc representations [7]. Secondly, a system's representations may or may not have some kind of intrinsic meaning. Those without require external interpretation for their representations to be understood. This is the case for classical symbol systems, where a theorist must provide an interpretation of its symbols for them to mean anything at all [8]. On the other hand, representations with intrinsic meaning are independent of being understood, similar to how a map can represent buildings and streets, regardless of whether it is taken as such [9].

These two characteristics — representational content and representational semantics — provide a way to classify cognitive systems based on how they represent the world. The traditional varieties of cognitive system implementations are re-cast along these lines in Figure 1. By having *unstructured content*, it is meant that the system's representations are either ad-hoc or non-existent, as opposed to those that are systematic. Representations with *external semantics* require meaning to be assigned by the designer, while those with *independent semantics* are exploitable by the system without intervention by a third party to provide meaning.

Content	Semantics	
	External	Independent
Unstructured	Dynamical Enactive	Connectionist
Systematic	Physical Symbol	?

Fig. 1. Cognitive system implementations classified by representational properties

Thus the traditional classification is replaced by one which classifies physical symbol systems as a kind with both systematic representations and external semantics, and while the various emergent systems all have unstructured representations, dynamical and enactive systems can be classified as having external semantics whereas connectionist systems have independent semantics. This not only better distinguishes between the different implementations, but significantly it also makes evident a missing kind of system — one with both systematic representations and independent semantics.

2 Symbolic Emergent Systems

The missing piece in the figure above poses a further question — what would a system with systematic, independent representations look like? Unlike existing hybrid approaches which simply attempt to combine differing implementations, the aim is for a system that makes use of representations with these specific properties, regardless of its implementation. By examining the features entailed by these properties, it is possible to put together a picture of such a system.

There is a strong argument that a symbol system is necessary for systematic representations. Systematicity requires both productivity — that an infinite number of possible representations can be formed using a finitely defined procedure, and compositionality — that complex representational content is constituted from other less complex representations, and the resulting combinatorial nature can only be met by a symbol system [7]. Further, for representations to be independently meaningful to a cognitive system, they must be in some way equivalent to the things they represent, in the same way as a map's streets are to those in the city. This implies representations are isomorphism of their targets — they have objects and relations which correspond to their target's objects and relations [9] and hence consist of some kind of relational structure. Thus the desired system should have a relational, symbolic basis for its representations[1].

What remains then is to explicate a kind of symbol system with representations that are exploitable without requiring a third party to provide meaning for them. One possible approach can be seen from the perspective of an agent built with this kind of system. Consider a hypothetical robot with only vision and temperature sensors, that has been newly situated in a typical kitchen. Over time it acquires visual percepts of an oven and thermodynamic percepts of hotness. Since it was not furnished a-priori with a symbolic theory of households, its representation of a "hot oven" must be composed of a relation over heat-percepts and representations of the oven, which is in turn composed of relations over visual percepts of same. Although a simplistic example, it is clear that the system's representational content can consist only of these things: percept symbols, other representations and relations over both. With these being the only objects of the system's causal machinery, what the system in effect requires is the instantiation of a symbolic theory of mind, rather than the usual classical approach of a collection of disparate theories of kitchens, thermodynamics, and so on.

This novel system, being a hybrid of the representational properties of both cognitivist and emergent systems, can perhaps be called a "symbolic emergent" system. It is a well-principled and yet unexplored foundation for cognitive systems, and is well-situated to be examined in a field such as cognitive robotics, where elements as disparate as symbol systems and laser sensors come together.

References

1. Nilsson, N.J.: Shakey the Robot. Technical report, SRI International (1984)
2. Newell, A., Simon, H.A.: Completer Science as Emprical Inquiry: Symbols and Search. Communications of the ACM 19(3), 113–126 (1976)
3. Brooks, R.A.: A Robust Layered Control System for a Mobile Robot. IEEE Journal of Robotics and Automation 2(1), 14–23 (1986)
4. Vernon, D., Metta, G., Sandini, G.: A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents. IEEE Trans. Evolutionary Computation 11(2), 151–180 (2007)
5. Anderson, M.L.: Embodied Cognition: A Field Guide. Artificial Intelligence 149(1), 91–130 (2003)
6. Brooks, R.A.: Intelligence without representation. Artificial Intelligence 47(13), 139–159 (1991)
7. Fodor, J.A., Pylyshyn, Z.W.: Connectionism and Cognitive Architecture: A Critical Analysis. Cognition 28(1-2), 3–71 (1988)
8. Slezak, P.: The Tripartite Model of Representation. Philosophical Psychology 15(3), 239–270 (2002)
9. Cummins, R.: Representations, Targets, and Attitudes. MIT Press (1996)

[1] Notably, while advocating a symbol system would usually invoke certain philosophical and technical problems such as the *Chinese room gedankenexperiment* and the *symbol grounding problem* they cease to be an issue when using representations with intrinsic meaning [9, Ch.9].