

# Extended Mind, Embedded AI, and “the Barrier of Meaning”

Sadeq Rahimi, MSc, PhD

Insync Strategy / Harvard Medical School / Boston Graduate School of Psychoanalysis  
[Sadeq.Rahimi@insyncstrategy.com](mailto:Sadeq.Rahimi@insyncstrategy.com)

## Abstract

This paper discusses notion of embeddedness as concerns the development of Artificial Intelligence, specifically contextualizing the notion in relation to theories of subjectivity developed in human and social sciences. It is argued that embeddedness not only coincides with significant themes of interest in the development of cognitive AI, such as insights from Behavioral Economics, or incorporation of more intricate social and cultural factors into the learning and cognitive algorithms, it can also be seamlessly converged with psychoanalytic and anthropological theories of subjectivity and symbolic systems of meaning towards conceptualization of subjective AI –artificial agents capable of understanding, producing, and communicating “meaning” in ways that “make sense” within human cultural and psychological systems.

## Extended Mind, Embedded AI, and “the Barrier of Meaning”

The mathematician and philosopher Gian-Carlo Rota’s famous question, “I wonder whether or when A.I. will ever crash the barrier of meaning,” (2008, p. 59) is often quoted, especially by those keen on conveying cynicism about AGI ever reaching the point of comparability to human intelligence. What is typically left out in those references, however, is the much more interesting “context” in which Rota’s question is expressed –a context that gives his question a very different “meaning” from what it appears to be asking at first. In the paragraphs leading to the famed question, Rota writes on behalf of Stan Ulam, “when you write down precise definitions for these words [words like keys, passenger, etc.] you discover that what you are describing is not an object, but a function, a role that is inextricably tied to some context. Take away the context, and the meaning also disappears” (pp 57-8). To make better sense, let’s put all this in the context of yet another famously quoted question with

which Clark and Chalmers open their seminal paper: “where does the mind stop and the rest of the world begin?” they ask (1998, p.7). Theirs is of course a rhetorical question, one to which their paper’s title, ‘The Extended Mind,’ has already given the answer. Now, to think extended mind is in fact to think embedded meaning –and those, extended mind and embedded meaning, I argue, are precisely what we need to be thinking as we think Artificial Intelligence.

As indicated in this year’s Symposium description, embeddedness requires closer and more deliberate attention as an aspect of AI, both in terms of how exactly we should understand the notion of embeddedness when speaking of AI (i.e. *conceptualization*), and in terms of what the practical relevance of embeddedness is for AI development (the *implications*). I will use this opportunity to focus primarily on the first aspect, and I will conclude with a brief overview of the implications of the ways we might understand and conceptualize embeddedness, specifically towards the understanding of subjective cognitive AI.

During early discussions of this year’s Symposium theme, a basic description of embeddedness as applied to AI was circulated as: “the dependence of an intelligent agent's activity on its environment, which may be defined alternatively in institutional, social, cognitive, or cultural terms.” Building on, and expanding this description, I would like to examine embeddedness as it pertains to theoretical and practical approaches to AI development, specifically concerning human subjectivity and the idea of subjective AI.

At the core of my argument here lies the basic assumption that subjectivity, human subjectivity that is, is an emergent phenomenon intrinsically contingent on the existence of a symbolic order. The notion of “symbolic order” is formulated here at an intersection of psychological anthropology (e.g. Victor Turner, 1974; Clifford Geertz, 1973) and structuralist French psychoanalysis (spearheaded by Jacques Lacan, 1968, etc.). Symbolic order denotes a structured semiotic environment that makes possible meaningful interactions of conscious subjects, and transmission of information through abstracted signs that stand for various internal (i.e. mental) and external objects and perceptions. It is a system of symbols, the units of which always function as signifiers, and the organization of which is closely tied to the organization of language and the organization of human mind. In fact, the connection is to the degree that, as Jacques Lacan has famously asserted, the human unconscious is structured as a language. And you can consider it both in terms of evolutionary history of human consciousness as a species (in the sense that the production and use of symbols coincides, or is in fact identical with, the emergence of consciousness); and in terms of the developmental trajectory of individual consciousness –In the sense that the human infant develops an ego, and a sense of the self as a point of reference that is distinct from the world around it, at precisely the same point in time when he or she becomes able to identify its own image in the mirror, and develops the capacity for conscious symbol making and mastery of language use.

What we call “culture” in this view, is nothing but an external memory device, a structure made of signs and symbolic relations, which is capable of storing information in abstracted form, so that for the first time (in the history of evolution) we become the species that is able to transmit information across generations through means other than genes, namely language and various other semiotic structures and devices, all of which can only function within a so-called symbolic order.

Drawing on these theories, I argue that, given the fundamental contingency of human subjectivity on the qualities of semiotic and material embeddedness, in order for an artificial agent to effectively interact with humans, it would need to be capable of a) synthesizing information to produce embedded meaning and b) processing and communicating such meaning through the sym-

bolic order to produce “subjective understanding” (cf. Carbonell, 1979).

Embedded meaning here of course refers to meaning that is by definition anchored and tied to the symbolic system that structures almost all aspects of our lives, ranging from the social order to language and other forms of communication, to politics, science, religion, economy and so on. Meaning is always embedded. In fact the way a sign comes to “mean” anything is by occupying a specific location in the system of negations and connections that constitutes our ordered experience of reality –look also back at Rota’s phrasing above, “not an object, but a function, a role that is inextricably tied to some context.” So in a sense “embedded meaning” is in fact a redundant phrasing. And as for “subjective understanding,” once again, when we understand the notion of subjectivity in terms of having a specific point of view in interpreting and perceiving an event or object, we are in fact making reference to a specific mode of orientation within the symbolic system which makes it possible to speak of a “point of view”. This so-called point of view is no longer defined merely in terms of physical/spatial and phenomenological relations it is now defined (and experienced) in terms of its reference and orientation within the system of signs, that symbolic system within which everything human is experienced, understood and communicated.

The notion of embeddedness was originally introduced to the field of economy by the Hungarian economic historian, Karl Polanyi (Polanyi, 1957) to capture the idea that the economic systems in “traditional” societies are fundamentally different from that in modern market societies, and that the difference can be understood as a matter of “embeddedness.”

According to Polanyi’s model, in so-called traditional societies an abstracted system of rational behavior as a structure that orders individual decision making has not yet emerged, and consequently, economic decisions are made based on a mixture of factors such as cultural values, social relationships, religious and moral convictions, political concerns, or the fear of authoritarian rulers. As a result, Polanyi’s theory says, the more abstract principles and formulations of economics cannot and should not be applied to those societies, because their decisions are not truly rational –they fall back on various interper-

sonal and irrational factors – so *that* is the sense in which economic systems in traditional societies is “embedded,” in the sense that it has not released itself to rise as an abstracted and abstract system. This quality, the theory goes, stands in contradistinction to modern market societies in which economy constitutes a separate and distinct (or dis-embedded, if you like) sphere, where exchange is regulated not by subjective processes, but in an abstracted vacuum governed by rules and regulations, rather than affects and relations.

It did not take long for social scientists including, eventually, economists themselves to realize the naiveté of this binary perspective. For instance, one of Polanyi’s own predecessors, Mark Granovetter, observed that even in so-called “modern” market societies economic processes are hardly abstract, fully rational and culture-free or dis-embedded processes. He was able to understand that economy, as a sphere of human decision making, is *always* embedded. The behaviors of economic actors in modern market societies too, he noted, are “embedded in concrete, ongoing systems of social relations” (Granovetter 1985, p. 487). And of course, needless to say, this conceptual development also went hand in hand with the other line of development in economic theories that we know as Behavioral Economics, and the notions of bounded rationality and the significant relevance of cognitive biases, which has been one of the core issues in our symposia –the awareness that homo-economicus as a fully rational agent can exist only in our imagination, not in real life.

“The challenge of creating humanlike intelligence in machines remains greatly underestimated,” writes Melanie Mitchell in a recent article, and she then goes on to say that, “today’s AI systems sorely lack the essence of human intelligence: [namely] *understanding* the situations we experience, [and] being able to grasp their meaning” (Mitchell, 2018). If economic theory has evolved to understand that the quality of embeddedness is not an undesirable quality associated with premodern life that market society has risen above, but rather a defining feature of human subjectivity and human cognition as such, then AI theory too needs to upgrade its conceptualization of successful AI from the dis-embedded abstraction of pure rationality to one of embedded subjectivity.

The core idea to understand about subjectivity as we work towards and think about seamless AI-human communication –and of course about the notions of AGI and artificial subjectivity, is that what we humans experience and identify as selfhood and subjectivity is neither created inside the single individual, nor limited to the individual’s physical boundaries. Both human consciousness and human subjectivity are collective phenomena –phenomena that would NOT exist as we know it in the absence of collective processes and collectively formed symbolic systems. They are contingent on the existence of a symbolic system within which they are processed and experienced, and they need the presence and interactivity of other subjects in order to bear (or to *make*) sense. This is not a new idea, this was one of the most significant contributions of Hegel’s philosophical edifice, his major book, *Phenomenology of the Spirit* is in many ways a detailed elaboration of this notion.

As even this brief discussion has clearly highlighted, a number of diverse lines of thought converge on the significance of thinking embeddedness in thinking about AI. The notion of embeddedness not only coincides with significant themes that are already paid attention to by AI researchers (such as cognitive biases and other insights from cognitive psychology and behavioral economics, or the increasing awareness of the need for inclusion of more intricate sociocultural factors in our thinking, research and development), but is also urged by major philosophical, psychoanalytic, and anthropological theories of selfhood and subjectivity. It would be hard to overemphasize the significance of understanding the idea that symbolic systems of meaning hold together our societies, communications, and the very sense of selfhood and subjectivity. It would be similarly difficult to overemphasize the need for conceptualization of artificial intelligence as an embedded agent capable of interpreting, producing, and communicating meaning in ways that “make sense” within those symbolic systems.

## References

- Carbonell, J. G. (1979). *Subjective Understanding: Computer Models of Belief Systems*. Doctoral Dissertation. New Haven: Yale University Department of Computer Science.
- Faucher, C. (2018). *Advances in Culturally-Aware Intelligent Systems and in Cross-Cultural Psychological Studies*. Gewerbestrasse, Switzerland: Springer.

Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of sociology*, 91(3), 481-510.

Hegel, G.W.F. *Phenomenology of Spirit*. Trans. A.V. Miller. New York: Oxford UP, 1977. Print.

Heimgärtner, R. (2018) Culturally-Aware HCI Systems. In C. Faucher (Ed.), *Advances in Culturally-Aware Intelligent Systems and in Cross-Cultural Psychological Studies*, pp. 11-37. Gewerbestrasse, Switzerland: Springer.

Lacan, J. (1968). *The Language of the Self: The Function of Language in Psychoanalysis*. Trans. Anthony Wilden. New York: Dell.

Mitchelle, M. (2018). Artificial Intelligence Hits the Barrier of Meaning. *The New York Times*, November 5, 2018.

Logan, R. K. (2007). *The extended mind: The emergence of language, the human mind, and culture*. University of Toronto Press.

Polanyi, K. (1957) The economy as instituted process. In K. Polanyi, C.M. Arensberg & H.W.

Pearson (Eds.) *Trade and Market in the Early Empires: Economies in History and Theory*. New York: The Free Press.

Rota, G. C. (2008). *Indiscrete thoughts*. Boston: Birkhauser.

Turner, V. (1974) *Dramas, fields and metaphors: Symbolic action in human society*. Ithaca: Cornell University Press.