Chapter 25
Tacit Representations and Artificial Intelligence: Hidden Lessons from an Embodied Perspective on Cognition

Elena Spitzer

Abstract In this paper, I explore how an embodied perspective on cognition might inform research on artificial intelligence. Many embodied cognition theorists object to the central role that representations play on the traditional view of cognition. Based on these objections, it may seem that the lesson from embodied cognition is that AI should abandon representation as a central component of intelligence. However, I argue that the lesson from embodied cognition is actually that AI research should shift its focus from how to utilize explicit representations to how to create and use tacit representations. To develop this suggestion, I provide an overview of the commitments of the classical view and distinguish three critiques of the role that representations play in that view. I provide further exploration and defense of Daniel Dennett's distinction between explicit and tacit representations. I argue that we should understand the embodied cognition approach using a framework that includes tacit representations. Given this perspective, I will explore some AI research areas that may be recommended by an embodied perspective on cognition.

Keywords Embodied cognition • Artificial intelligence • Representation • Tacit representations

25.1 Introduction

In the past 20 years, embodied cognition (EC) has emerged as a proposed alternative or enhancement to traditional cognitive science. Embodied cognition can be seen as a research program that encourages us to pay more attention to the role that the rest of the body, not just the brain, plays in cognition. One common EC objection against traditional cognitive science has to do with the role that internal representations are presumed to play in cognition. When we pay more careful attention to the role of

E. Spitzer (✉)
University of Wisconsin, Madison, WI, USA
e-mail: emspitzer1@gmail.com

V.C. Müller (ed.), *Fundamental Issues of Artificial Intelligence*, Synthese Library 376, DOI 10.1007/978-3-319-26485-1_25

the body in cognition, it often seems that the types of representations postulated by traditional cognitive science are explanatorily inadequate.

In this essay, I consider what lessons an embodied perspective on cognition might hold for the development of artificial intelligence. Artificial intelligence had ambitious beginnings in the latter part of the twentieth century, with grand predictions of human–level intelligence right around the corner. However, AI research has since run into a number of difficult challenges, and the original AI dream of general intelligence has faded. Current research on and applications of AI (especially the most successful ones) focus on "specialized intelligence." In other words, the focus is on ways to make computers better at the sorts of tasks that humans are bad at—such as massive information storage and fast retrieval, very large scale data analysis, matching advertisers with search engine users, etc. There is less focus on "general intelligence"—how to build a computer that demonstrates some of the sorts of intelligence that come so naturally to humans. Progress towards building a computer that is able to do tasks that a five year old could perform easily, like playing hide-and-go-seek, has been slow.

In this paper, I will explore how an embodied perspective on cognition might inform research on developing general intelligence in AI. I will discuss arguments from three embodied cognition theorists—Andy Clark, Hubert Dreyfus, and Rodney Brooks. As critics of the classical view of cognition, each of these authors objects to the central role that representations play on the traditional view. Based on these arguments, it may seem as though the lesson from embodied cognition is that AI should abandon representation as a central component of intelligence.

However, I will argue that the lesson from embodied cognition is actually that AI research should shift its focus from how to utilize *explicit* representations to how to create and use *tacit* representations. In Sect. 25.1, I will provide an overview of the commitments of the classical view of cognition. Section 25.2 describes the role of representations in the classical view and three different critiques of that role. In Sect. 25.3, I introduce Daniel Dennett's distinction between three types of representation (explicit, implicit, and tacit), and provide further exploration and defense of tacit representations. The final section demonstrates how the embodied cognition perspective from Clark, Dreyfus, and Brooks could be helpfully reconceived in a framework that includes tacit representations. In this final section I will also explore some AI research areas that may be recommended by an embodied perspective on cognition.

25.2 The Classical View

The standard view in cognitive science is that cognition consists of a certain type of computational process. More specifically, the brain is presumed to be some kind of symbol system, and cognition proceeds by symbol manipulation. Newell and Simon state the Physical Symbol System Hypothesis (PSSH) as follows:

> *The Physical Symbol System Hypothesis*. A physical symbol system has the necessary and sufficient means for general intelligent action. (Newell and Simon 1976, p. 292)

The PSSH is taken to be an empirical hypothesis, and some version of this hypothesis has formed the basis of most theorizing and research in cognitive science. On the basis of the PSSH, cognitive science set out to locate the cognitive symbols and the rules that govern their manipulation.

The idea that the brain is a physical symbol system generally involves (at least) the following commitments. First, physical parts of the brain act as symbols in the cognitive system. For something to be a symbol means that it represents something else (i.e. has some semantic content), that its relationship to what it represents is arbitrary (e.g. the word "dog" does not look like a dog), and that it can potentially be combined with other symbols to create new symbols (Shapiro 2010, pp. 10–11).

Additionally, the PSSH includes the notion that thinking involves the process of symbol manipulation. For example, the cognitive process of reasoning from the proposition "Socrates is short and Socrates is ugly" to the proposition "Socrates is short" is analogous to the standard inference rule of conjunction elimination in first-order logic. There is some part of the brain that serves as a symbol for "Socrates is short" and another symbol for "Socrates is ugly." In sentential logic, we derive "S" from "S & U" based on the syntactic structure of the conjunction of the two symbols "S" and "U" and the rule of conjunction elimination. Analogously, according to the PSSH, the brain reasons from "Socrates is short and Socrates is ugly" to "Socrates is short" according to the syntactic structure of the proposition and rules about reasoning.

One particularly influential elaboration on the PSSH is Jerry Fodor's Language of Thought Hypothesis. The Stanford Encyclopedia describes this hypothesis as follows:

> The Language of Thought Hypothesis (LOTH) postulates that thought and thinking take place in a mental language. This language consists of a system of representations that is physically realized in the brain of thinkers and has a combinatorial syntax (and semantics) such that operations on representations are causally sensitive only to the syntactic properties of representations. According to LOTH, thought is, roughly, the tokening of a representation that has a syntactic (constituent) structure with an appropriate semantics. Thinking thus consists in syntactic operations defined over such representations. (Aydede 2010)

The standard view holds that cognition proceeds via the algorithmic manipulation of syntactically structured explicit representations in the brain.

25.3 Embodied Cognition Critiques

A number of criticisms of the classical view in cognitive science and its attendant research programs in AI have focused on the role that representations are presumed to play in cognition—human or artificial.

I will discuss three different proponents of embodied cognition who have criticized the role of representations in the classical view in one way or another—Andy Clark, Hubert Dreyfus, and Rodney Brooks. Broadly speaking, embodied cognition theorists argue that the classical view in cognitive science is *too cerebral* in some sense. Specifically, they think that the classical view does not take sufficient account of the role that the body (other than the brain) plays in cognition.

The classical view presumes that the body's role in cognition is fairly limited. The body serves as an input and output device—delivering signals to the brain and carrying out its commands. It is supposed that the truly cognitive processes consist only of what happens between the input and output. Shapiro puts the point as follows: "All the 'action,' so to speak, begins and ends where the computational processes touch the world. The cause of the inputs and the effect on the world that the outputs produce are, from this perspective, irrelevant for purposes of understanding the computational processes taking place in the space between input and output" (Shapiro 2010, p. 26).

Against the classical view, proponents of embodied cognition claim that the body often plays an important role in cognition, which the classical view fails to recognize. Thus, EC proponents maintain that cognitive science must take better account of cognitive agents' bodies in order to understand cognition more fully. The evidence for this line of thinking is varied, but the message is similar.

25.3.1 The Role of Representations

Representations play a particularly important role in the classical view. For example, Fodor and Pylyshyn characterize the classical view in terms of (1) the properties of mental representations and (2) the properties of the processes that transform those mental representations (Fodor and Pylyshyn 1988, pp. 12–13). Representations are the symbols part of a physical symbol system. And they are central to the Language of Thought Hypothesis, according to which internal representations are the components that make up the language of thought.

Because of this prominent role, representations are often a target for objections to the classical view. Arguments against the classical view from proponents of embodied cognition often follow this general pattern: The body plays such-and-such a role in cognition. Representations are the only means of explaining cognitive activity on the classical view. But representations cannot explain the role that the body plays in so-and-so. Therefore, the classical view is insufficient for explaining cognition.

Below, I will discuss three arguments from EC proponents that target the role of representations in the classical view. Each of the arguments finds that representations are inadequate for various reasons, and thus concludes that the classical view must be amended, augmented, or perhaps even abandoned.

25.3.2 Clark's Cricket

Internal representations are a necessary component of cognition on the classical view. However, they often seem ill-suited to the more embedded and embodied components of cognition that EC theorists focus on. Some EC theorists advocate abandoning internal representations entirely. Clark does not agree that we can do away with internal representations altogether, but he does think it may be appropriate to deemphasize internal representations in many areas of cognition.

Clark discusses the following example from Barbara Webb's research as a case in which a seemingly intellectual and representation-dependent task is better understood as a mechanistic and non-symbolic task.

> Phonotaxis is the process whereby a female cricket identifies a male of the same species by his song, turns in his direction, and reliably locomotes to the source. On the face of it, this looks to be a complex, three part problem. First, hear various songs and identify the specific song of the male of your own species. Second, localize the source of the song. Third, locomote that way (making needed corrections en route). This way of posing the problem also makes it look as if what the cricket will require to solve the problem are some quite general cognitive capacities (to recognize a sound, to discern it's source, to plot a route). Nature, however is much thriftier, and has gifted the cricket with a single, efficient, but totally special-purpose problem-solving procedure . . .
>
> This whole system works only because first, the cricket's tracheal tube is especially designed to transmit songs of the species-specific frequency, and because the male song repeats, offering frequent bursts each of which powers one of these episodes of orienting-and-locomoting (hence allowing course corrections, and longer distance mate-finding) . . .
>
> In the very specific environmental and bio-mechanical matrix that Webb describes, the cricket thus solves the whole (apparently three-part) problem using a single special–purpose system. There is no need, for example, to . . . bother to build a model of your local surroundings so as to plan a route. Instead, you (the cricket) exploit neat tricks, heuristics, and features of your body and world. Moreover you (and your realworld sisters) seem to succeed *without relying on anything really worth calling internal representations*. (Clark 2001, pp. 127–128 emphasis added)

There is no agreed upon definition of what makes an activity cognitive. But for the sake of argument, I think we can treat phonotaxis as cognitive here. Though I think there are other good reasons to count phonotaxis as cognitive. One reason is that this type of flexible interaction with the world seems to be essential for general intelligence. And, insofar as my project here concerns general artificial intelligence, we need some account of how this type of activity works. Especially since this is exactly the kind of thing that computers are much worse at than living organisms. Rodney Brooks, a roboticist that we'll meet again later in the paper, stresses the importance of this type of "low-level" activity: "Our goal . . . is simple insect level intelligence within two years. Evolution took 3 billion years to get from single cells to insects, and only another 500 million years from there to humans. This statement is not intended as a prediction of our future performance, but rather to indicate the nontrivial nature of insect level intelligence" (Brooks 1991, p. 156). Being able to explain activities such as phonotaxis is likely an important step in better understanding cognition and developing general AI. So, even if we do not want to call phonotaxis cognitive *per se*, it does seem to be the type of activity that is a precondition for "real" cognitive activity—and that's cognitive enough for my purposes here.

According to the classical view, the most natural way to break down the cricket's behavior is as a three step process—recognize, locate, and move towards the song. On this model, the female cricket has internal representations of what she hears, of whether it matches the song in which she is interested, of where the sound is coming from, how to get there, etc. But the way that phonotaxis actually happens does not seem to mirror this structure at all. The observed methods seem to have less in common with rule-based inference and more in common with simply mechanical operations—no rules and no representations, according to Clark.

Clark suggests that the sort of mechanisms responsible for cricket phonotaxis may be very common cognitive mechanisms in humans as well. But this requires that the classical view either be supplemented or replaced.

25.3.3 Dreyfus & Dreyfus and the Foundations of AI

Another philosopher who has criticized the role of representations in artificial intelligence is Hubert Dreyfus. Hubert Dreyfus and his brother, Stuart Dreyfus, provided an early and influential critique of artificial intelligence.

Representations play a central role in the Dreyfus & Dreyfus critique of the classical view. The quick, but hopefully not too dirty version of the argument goes like this: AI is doomed because genuine intelligence requires a commonsense background that computers can never possess. Dreyfus says, "Intelligence requires understanding, and understanding requires giving a computer the background of common sense that adult human beings have by virtue of having bodies, interacting skillfully with the material world, and being trained into a culture" (Dreyfus 1992, p. 3). The type of commonsense background that Dreyfus has in mind can be demonstrated by his example of the "cultural *savoir faire*" of gift-giving—the ease with which an adult human in a particular culture can give an appropriate gift at an appropriate time and in an appropriate way (Dreyfus 1992, p. xxiii). For example, when you are walking through a shop looking for a gift to take to a friend's dinner party, you may stop to consider a bouquet of flowers, but you do not stop to consider bringing 20 cans of tuna fish or a box of roach motels.

Intelligence requires a commonsense background, but Dreyfus & Dreyfus think this background cannot be represented. They believe that this background consists of a set of skills, not a body of knowledge—i.e. commonsense background requires knowledge-how, not knowledge-that. However, Dreyfus & Dreyfus believe this is the sort of information that cannot be represented by computers. They say, "If background understanding is indeed a skill and if skills are based on whole patterns and not on rules, we would expect symbolic representations to fail to capture our commonsense understanding" (Dreyfus and Dreyfus 1988, p. 33). Dreyfus & Dreyfus think that the commonsense background cannot be enumerated and represented as a set of facts, and that this is what would be required in order to get a computer to act intelligently. Thus, because of their reliance on representations, the classical view will be incapable of fully explaining intelligence and AI will be incapable of building genuine intelligence.

25.3.4 Brooks' Creatures

Rodney Brooks is a roboticist from MIT. In his article "Intelligence Without Representation," he describes some of his work building intelligent autonomous mobile robots, which he calls "creatures." Brooks' creatures were groundbreaking because they were relatively good at performing tasks in real-world environments. Previous robots could perform in highly structured environments, but could not deal with the variability of a real world environment.

Brooks attributes his creatures' success in real-world environments to his unique approach to their design—an approach that does not use representations as traditionally conceived. Previous approaches to mobile robot design had the robot create representations of the world based on information gained from its sensors. These robots would then have a central processor reason about those representations in conjunction with its goals (also represented) in order to decide what to do next.

Brooks calls his design approach the "subsumption architecture." On this approach, instead of having external sensors and a centralized internal processor, there are multiple layers of control, each of which interacts with the world and "makes decisions" on its own.

Like Clark and the Dreyfus brothers, Brooks also wants to diminish or eliminate the role of representation in cognition. Indeed, the title of his 1991 paper is "Intelligence Without Representation." Brooks attributes the success of his creatures to the absence of representations. He says, "a common theme in the ways in which our layered and distributed approach helps our Creatures meet our goals is that there is no central representation" (Brooks 1991, p. 147).

Brooks' work demonstrates that "there need be no explicit representation of either the world or the intentions of the system to generate intelligent behaviors for a Creature" (Brooks 1991, p. 149). This does not show that intelligent behavior never requires explicit representations, but it certainly seems to demonstrate a more promising approach to realworld interaction than previous attempts. And, if Brooks is right that being able to act in a complex and dynamic environment is a cornerstone of intelligence, then this is potentially a more promising approach for AI than the one more informed by the classical view.

25.3.5 Ostensible Lesson for AI

All of these theorists take issue with the role that representations have played in standard accounts of cognition. Clark argues that the classical view is too quick to posit representations, when apparently non-representational mechanisms might account for cognitive phenomena. Dreyfus and Dreyfus argue that genuine intelligence requires knowhow and that this knowhow cannot be represented. For them, this implies that computers are not capable of intelligence because computers are limited to using information that can be represented. Finally, Brooks argues that representations are the wrong way to proceed in order to build robots that can interact effectively with a dynamic environment. He eschews representations in favor of an alternative design strategy that he claims does not rely on representations at all.

All three viewpoints seem to suggest that the lesson for AI from an embodied perspective on cognition is that if AI wants to achieve general intelligence, then AI theorists should shift their focus away from representations. The lesson seems to be that representations are the wrong way to go in order to achieve general intelligence of the sort displayed by humans (or other animals). In varying ways and to varying degrees, all three authors suggest that representations are inadequate to account for the kind of skilled interaction that is necessary for general intelligence.

This *seems* to be the lesson for AI from an embodied perspective on cognition, but I think it is problematic. I think this lesson is problematic because any intelligent system must have some way of obtaining and using information about its environment. And anything that serves the role of conveying information is a representation of some sort. So obtaining and using information about one's environment actually requires representations. Thus, advising AI theorists to abandon representations is incoherent if it amounts to a suggestion that they should give up trying to build artificial systems that obtain and use information about the environment. I will revisit this idea again below.

I will argue that there is actually a different lesson for AI from an embodied perspective on cognition. Instead of advising AI to abandon representations, I think the lesson is that AI theorists should focus on a particular kind of representation—tacit representations. I will explain what tacit representations are and how they relate to embodied cognition and AI in the next two sections.

25.4 Are All Representations Created Equal?

Each author discussed above thinks that at least some cognition does not involve representations at all. Their work also suggests the even stronger claim that non-representational cognitive mechanisms are a necessary feature of intelligence. Below I will argue that we can develop an alternative to the classical view without giving up on representations altogether.

25.4.1 Basics of Representation

In order to proceed, we first need to have some working definition of representation. There is no agreed upon definition, but Fred Dretske has developed a plausible account, and I'll use that as my working model here. Prinz and Barsalou (2000) adopt this approach as well and provide the following useful summary of Dretske's views:

> Many philosophers believe that representation involves information. A state represents a property only if it carries information about it (Dretske 1981; Fodor 1990). Carrying information is, in turn, analyzed in terms of nomic, or law-like, covariation. To a first approximation, a state s carries information about a property F just in case instantiations of F reliably cause tokens of s. Although arguably necessary, information is not sufficient for representation. Fire reliably causes smoke, but smoke does not represent fire. An information-bearing state must satisfy one further condition to count as a representation. For Dretske (1995), this extra ingredient is teleological: Something can represent something else only if it has the function of carrying information about it. Representations can acquire such functions through a variety of different histories, including natural selection, learning, and design. (Prinz and Barsalou 2000, p. 55)

So, there are two components of representations on Dretske's account—an information component and a teleological component. Below, I will outline three types of representation and use Dretske's account to assess one of them.

25.4.2 Three Styles of Representation

In "Styles of Mental Representation," Daniel Dennett distinguishes between three types of representation: explicit, implicit, and tacit. Dennett wants to explore the ways in which computers represent in order to investigate the analogous features in human cognition.

Dennett suggests that cognitive science has been overly focused on explicit representations. Explicit representations are the most obvious kind of representations in a computer because these are what we deal with when we program computers. Perhaps as a result of this, the classical approach to cognitive science has focused on analogous theoretical structures in the human mind. However, Dennett argues that computers also contain tacit representations, and that these might be more informative and useful as a theoretical tool in cognitive science than has previously been appreciated.

Explicit representations are the most familiar type of representation. The word "cat" explicitly represents a cat, the word "bicycle" explicitly represents a bicycle, and the picture below (Herford 1899) explicitly represents a cat riding a bicycle.

Each of these is an explicit representation because it is a physical object that has semantic content according to some system of interpretation. An explicit representation is physically stored in a system and can be interpreted and used by the system (Dennett 1982).

Dennett defines an implicit representation as one that is logically implied by explicit representations in the system. For example, if a system explicitly represents "Eeyore is a donkey" and "All donkeys have ears," then it implicitly represents "Eeyore has ears." Thus, you could (implicitly) represent the proposition "Eeyore has ears," even if you had never considered the matter of Eeyore's ears before.

Dennett distinguishes a third type of representation—tacit representations. Dennett's description of tacit representations is less well–developed than his description of either explicit or implicit representations. He uses the example of a pocket calculator to clarify the idea. The example starts by asking, "Does a pocket calculator represent the 'truths of arithmetic' explicitly, implicitly, or tacitly?" (Dennett 1982, p. 221).

The calculator produces correct answers to arithmetical queries without explicitly referencing any rules of arithmetic. When the calculator computes "$2+2$," it does not produce "4" by looking up a rule that says "$2+2=4$." Dennett says, "the calculator is a device with the dispositional competence to produce explicit answers to explicit questions … but it does this without relying on any explicit representations within it—except the representations of the questions and answers that occur at its input and output edges and a variety of interim results" (Dennett 1982, p. 222). The calculator's circuitry is designed such that it performs addition according to the rules of arithmetic, but these rules are not themselves explicitly represented in the system. The rule that "$2+2=4$" is not stored in the system and recalled and used when necessary. Nor are the rules of arithmetic implicitly represented by the calculator. Instead, engineers built the calculator so that it would operate in a way that conforms to the rules of arithmetic. Thus, we might say that the calculator tacitly represents how to do arithmetic.

As a first attempt to generalize from this example, perhaps we can say that a representation is tacit when it is, in some sense, *built into* a system. An explicit representation is something that the system *uses as* a representation in order to accomplish some task. For example, if we had a calculator that worked by using a lookup table, then the entries in the look up table would be explicit representations because of how they are used by the system. By contrast, although a tacit representation also has a physical presence in the system, it is not used *by* the system. Instead, a tacit representation is the thing in virtue of which the system displays some disposition.

The calculator's circuitry makes it possible for the calculator to do addition—to have the disposition to produce the correct input-output pairs. But the calculator does not consult its circuitry in order to determine how to perform addition. Instead, the calculator's particular circuitry is the mechanism in virtue of which it is able to perform addition. The rules of arithmetic are not stored and consulted. Nor does the circuitry itself consult any rules. The rules of arithmetic are the reason that the circuitry does what it does—since they guided its design. The calculator has arithmetic knowhow because its circuitry tacitly represents how to do arithmetic.

25.4.3 A Case for Tacit Representations as Representations

Some might object that knowhow is not the kind of thing that can be represented. Propositions can be represented but skills and dispositions cannot. However, I think there are good reasons to consider tacit representations as *bona fide* representations.

According to Dretske's account, representations comprise both an information component and a teleological component. Below I will argue that tacit representations include both components.

The information criterion says that state *S* represents *Q* only if *S* carries information about *Q*. In the case of tacit representations, I propose that a system that tacitly represents how to Q contains information about how to Q. The nature information is a debated topic (Adriaans 2012; Floridi 2013), and I do not propose to attempt to settle the matter here. And so, in order to make use of Dretske's information criterion, I will appeal to an everyday (though admittedly vague) understanding of the word.[1]

[1]This is Dretske's strategy as well. For example, in "The Epistemology of Belief," Dretske says, "I have begun to talk more and more about information, so let me pause a moment to explain what I mean by this way of talking. I mean nothing very technical or abstract. In fact, I mean pretty much what (I think) we all mean in talking of some event, signal or structure carrying (or embodying) information about another state of affairs. A message (i.e., some event, stimulus or signal) carries

On this everyday understanding, I think it is reasonable to say that knowhow involves using information about how to do something. For example, when someone knows how to do a flip-turn while swimming, they have information about how to do a flip-turn. One must learn how to do a flip-turn, and I think this can plausibly be understood (at least in part) as gaining information about how to do one. Knowhow requires information, and I propose that tacit representations carry that information.

According to the discussion above, a tacit representation is the structure(s) in virtue of which a system performs according to its knowhow. Given this definition of a tacit representations and the idea that knowhow has informational content, I think we must recognize tacit representations as carrying information about knowhow. The pocket calculator is able to do arithmetic, and its circuitry contains information about how to do arithmetic. My "circuitry" contains information about how to do a flip-turn, and so I know how to do a flip-turn. Thus, tacit representations meet Dretske's information criterion.

Dretske's second criterion is the teleological criterion. The teleological criterion requires that representations *have the function* of representing the thing which they represent.

I think tacit representations meet this criterion as well. Things acquire functions in different ways, so we may need to judge on a case-by-case basis. But there is nothing about tacit representations in and of themselves that would seem to violate this condition. And certainly in the case of the calculator's circuitry, by design, it has the function of carrying information about how to do arithmetic.

Thus, it is not the case that knowledge-that can be represented while knowledge-how cannot. Both can be represented in a system. But the role that each plays in the system is quite different. Knowledge-that is explicitly or implicitly represented, while knowledge-how is tacitly represented.

25.5 Lessons for AI

25.5.1 A False Dilemma

Embodied cognition proponents often seem to suggest that we must choose between representation and embodied cognition. However, if we employ a framework that includes tacit representations, then this is turns out to be a false dilemma. I think we need not choose between representations and an embodied approach to cognition.

The classical view says that cognition involves the manipulation of explicit representations on the basis of syntactic rules. All three of the embodied cognition theorists discussed above reject this view because it cannot account for various important cognitive phenomena. The notion of internal representations is clearly an

information about X to the extent to which one could learn (come to know) something about X from the message" (Dretske 1983, p. 10).

important component of the classical view. And, I think these embodied cognition theorists are right to contend that the role representations play in the classical view is particularly ill-suited to account for certain important cognitive phenomena. But that does not mean that representations themselves are not a part of the explanation for these phenomena. Instead, it only implicates the particular commitments that the classical view has concerning representations—namely that explicit representations should be used to explain cognitive phenomena.

But tacit representations are not faulted in the same way as explicit representations by these arguments. In fact, I think a framework that includes tacit representations can more easily accommodate the examples that Clark, Brooks, and Dreyfus discuss. I will revisit each of these author's examples below to demonstrate how the examples would fit into the tacit representation framework.

25.5.2 Crickets, Creatures, and Commonsense

25.5.2.1 Crickets

Clark describes the crickets' phonotaxis machinery as a case where there is no representation of the world or the rules that guide the crickets' behavior. Instead, he says, the crickets "seem to succeed without relying on anything really worth calling internal representations" (Clark 2001, p. 128). Clark does not tell us what makes something worthy or unworthy of the name "internal representation," but I think it's safe to assume he has explicit representations in mind.

I think Clark is right to judge that explicit representations are out of place in explaining cricket phonotaxis. However, I think it would be quite reasonable to understand the cricket's phonotaxis machinery as a tacit representation. Features of the cricket's bodily setup tacitly represent its mate-detection-knowhow. There are a lot of details included in this knowhow—how to recognize a potential mate's song, how to locate him, how to get there, and a multitude of other elements. But, as Clark argues, it is the cricket's physical features in virtue of which it is able to perform this task. This seems to fit quite well into the tacit representation framework discussed above. Unlike the calculator's circuitry, the cricket's circuitry was not *designed* to function in this manner. But it sill has an appropriate teleology—since presumably it was *selected for* in order to perform this function.

25.5.2.2 Creatures

Brooks offers a sentiment that is similar to Clark's about what's "worth" calling a representation:

> Even at a local level we do not have traditional AI representations. We never use tokens which have any semantics that can be attached to them ... An extremist might say that we really do have representations, but that they are just implicit ... *However we are not happy*

> *with calling such things a representation. They differ from standard representations in too many ways*... There are no choices to be made. To a large extent the state of the world determines the action of the Creature. (Brooks 1991, p. 149, emphasis added)

Brooks occasionally refers to the representations he's rejecting as "explicit representations," but he's not consistent in his usage. He often seems to suggest that his creatures do not use any representations at all.

I think a representational framework that includes tacit representations is compatible with and even complements Brooks' subsumption architecture idea. Within this framework, Brooks' focus on system architecture design can be understood as a technique to implement knowhow in a system. The subsumption architecture enables the creatures to perform skillfully in their environment without having to represent facts about the environment and reason about them.

Brooks suggests that the fact that the state of the world (directly) determines the action of his creatures is a reason in favor of rejecting the idea that there are representations involved. Instead of representing the world and then using rules to reason about those representations and decide on an action, Brooks' creatures are designed so that they interact with the world in a more direct fashion. This is telling against explicit representations, but not so for tacit representations. Instead, I think tacit representations are quite suitable for describing Brooks' subsumption architecture.

25.5.2.3 Commonsense

In *What Computers Still Can't Do*, Hubert Dreyfus elaborates on his objection to AI and representations. He discusses the rationalist history of AI—philosophers such as Descartes and Leibniz thought the mind could be defined by its capacity to form representations of the "fixed" and "context-free" features of a domain and the rules governing their interactions. According to Dreyfus, on this view everything that we know, including knowhow, must be stored in the mind in propositional form. He calls this view of the mind "representationalism" (Dreyfus 1992, p. xvii).

Dreyfus argues that genuine intelligence requires understanding and understanding requires a rich background that can only be obtained by having a body and interacting constantly and skillfully with the material and cultural world (Dreyfus 1992, p. 3). With representationalism as the foundation, traditional AI assumes that "common sense derives from *a vast data base of propositional knowledge*" (Dreyfus 1992, p. xvii). The "common sense problem" for AI lies in how to go about creating, structuring, and querying this database so as to create genuinely intelligent action. Dreyfus says, "All attempts to solve [these problems] have run into unexpected difficulties, and this in turn suggests that there may well be in-principle limitations on representationalism" (Dreyfus 1992, p. xviii).

Despite the broad term "representationalism," I think Dreyfus is primarily concerned with the failings of explicit representations in particular. This is suggested by his focus on propositional content in his discussion of representations. But, if this is right, then tacit representations are not necessarily implicated by his arguments. Dreyfus objects to the “classical” version of AI on the grounds that commonsense and skillful interaction are necessary components of intelligence. But, he argues, commonsense and skillful interaction cannot simply be enumerated, represented, and effectively used by a system.

However, this same enumeration strategy is not required for a system to tacitly represent something. So the insufficiency of enumeration and explicit representation does not necessarily undermine tacit representations in the same way.

I think that adopting a framework that includes tacit representations, rather than just explicit representations, suggests a more promising strategy for meeting Dreyfus’s commonsense challenge. Though, of course, many of the details still need to be worked out. But this just means that additional research is required, not that these commonsense problems are insoluble within a framework that includes representations.

25.5.3 A Hidden Lesson

Despite the oft professed repudiation of representation from EC proponents, I think it is more fruitful to understand the lessons for cognitive science and artificial intelligence in a more nuanced way. Our takeaway from EC proponents’ discussion of representation should be that *explicit representations* are not up to the task of accounting for some of the foundational components of cognition or intelligence.

The focus of AI work on representation has generally been at a relatively high level of abstraction—e.g. building internal models of the world, using explicit rules, and symbolic logical reasoning. This mirrors the theoretical framework and focus of research in cognitive science. I think we should interpret results from work on embodied cognition as suggesting that cognitive science and AI should shift more focus onto understanding the value of tacit representations for cognition.

One negative project in embodied cognition points out that the classical view, with its focus on explicit representations, cannot account for important cognitive phenomena. But there is a positive project as well. This positive project suggest that important components of cognition are more mechanistic or disposition-based than previously appreciated. I think this positive project actually suggests adopting a framework that includes tacit representations rather than abandoning representations as a way to explain these sorts of cognitive phenomena.

An embodied perspective on cognition should not be used to admonish AI theorists to abandon representations—this is impractical at best and incoherent at worst. Intelligence requires that a system have a way of obtaining and using information about the environment. Artificial intelligence must be designed, so AI engineers must build mechanisms that allow a system to obtain and use information about the environment. But, at some basic level, this is all that is meant by representation—whatever serves the role of conveying information. Advising AI theorists to abandon representations is incoherent if it amounts to a suggestion that they should give up trying to build artificial systems that obtain and use information about the environment. It is impractical if it amounts to a suggestion that AI theorists should start from scratch in figuring out how to build a system that can obtain and use information.

Instead, the lesson for AI from an embodied perspective on cognition might be more like this: Given what we know about human and animal intelligence and cognition, an important (though previously underappreciated) component for intelligent action seems to be knowledge that is tacitly represented by the system. Work in embodied cognition suggests that some knowledge must be built into the architecture of a system so that the system can display appropriate dispositions.

What I mean by system architecture should be interpreted rather broadly here—both the calculator's circuitry and Brook's subsumption architecture would qualify. Although couched in different terms, connectionist architectures would also fit under this umbrella. Thus, I think one research area embodied cognition suggests is to further investigate the landscape of system architecture within a framework that includes tacit representations. What design principles should guide the design of system architectures such that they are endowed with those tacit representations required for intelligence? How can we integrate system architecture into cognitive tasks?

Another research area that seems especially interesting is how to design systems that can usefully modify their architecture. This ability seems to account for much of the intelligence we find in the animal world. Certainly evolution has resulted in physical architectures that convey important information, allowing organisms to act intelligently. And in the case of humans, our intelligence is often attributed to our comparatively abundant neural plasticity. It would be interesting to consider how to develop analogous abilities for artificial systems.

Continuing in the evolutionary direction, and taking a cue from Brooks' subsumption architecture, it would also be interesting to investigate the foundations and history of cognition in the natural world. Can we identify particular skills that are foundational for intelligent behavior? Is there any particular order in which these skills should be developed or implemented in a system? Studying the evolution of cognition with the tacit representation framework in mind should provide many ideas helpful to developing artificial intelligence that can mimic natural intelligence.

References

Adriaans, P. (2012). Information. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2012.). Retrieved from http://plato.stanford.edu/archives/win2012/entries/information/

Aydede, M. (2010). The language of thought hypothesis. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010). Retrieved from http://plato.stanford.edu/archives/fall2010/entries/language-thought/

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence, 47*(1), 139–159.

Clark, A. (2001). Reasons, robots and the extended mind. *Mind & Language, 16*(2), 121–145.

Dennett, D. C. (1982). Styles of mental representation. *Proceedings of the Aristotelian Society, 83*, 213–226. doi:10.2307/4545000.

Dretske, F. I. (1983). The epistemology of belief. *Synthese, 55*(1), 3–19. doi:10.2307/20115855.

Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. Cambridge, MA: MIT Press.

Dreyfus, H. L., & Dreyfus, S. E. (1988). Making a mind versus modeling the brain: Artificial intelligence back at a branchpoint. *Daedalus, 117*, 15–43.

Floridi, L. (2013). Semantic conceptions of information. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2013.). Retrieved from http://plato.stanford.edu/archives/spr2013/entries/information-semantic/

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1–2), 3–71. doi:10.1016/0010-0277(88)90031-5.

Herford, O. (1899). Retrieved from http://commons.wikimedia.org/wiki/File:Kitty_Riding_Bicycle.jpg

Prinz, J. J., & Barsalou, L. W. (2000). Steering a course for embodied representation. *Cognitive dynamics: Conceptual change in humans and machines*, 51–77. Cambridge, MA: MIT Press.

Shapiro, L. (2010). *Embodied cognition*. Routledge: Taylor & Francis.