Embodiment: Does a laptop have a body?

Pei Wang
Temple University, Philadelphia, USA
http://www.cis.temple.edu/~pwang/

Abstract

This paper analyzes the different understandings of “embodiment”. It argues that the issue is not on the hardware a system is implemented in (that is, robot or conventional computer), but on the relation between the system and its working environment. Using an AGI system NARS as an example, the paper shows that the problem of disembodiment can be solved in a symbolic system implemented in a conventional computer, as far as the system makes realistic assumptions about the environment, and adapts to its experience.

This paper starts by briefly summarizing the appeal for embodiment, then it analyzes the related concepts, identifies some misconceptions, and suggests a solution, in the context of AGI research.

The Appeal for Embodiment

In the last two decades, there have been repeated appeals for embodiment, both in AI (Brooks, 1991a; Brooks, 1991b; Pfeifer and Scheier, 1999) and CogSci (Barsalou, 1999; Lakoff and Johnson, 1998). In AI, this movement argues that many problems in the field can be solved if people move their working platform from conventional computer to robot; in CogSci, this movement argues that human cognition is deeply based on human sensorimotor mechanism.

In general, “embodiment” calls people’s attention to the “body” of the system, though like all theoretical concepts, the notion of “embodiment” has many different interpretations and usages. This paper does not attempt to provide a survey to the field, which can be found in (Anderson, 2003), but to concentrate on the central issue of the debate, as well as its relevance to AGI research.

The stress on the importance of body clearly distinguishes this new movement from the traditions in AI and CogSci. In its history of half a century, a large part of AI research has been guided by the “Physical Symbol Hypothesis” (Newell and Simon, 1976), which asks AI systems to build internal representation of the environment, by using “symbols” to represent objects and relations in the outside world. Various formal operations, typically searching and reasoning, can be carried out on such a symbolic representation, so as to solve the corresponding problems in the world. Representative projects of this tradition include GPS (Newell and Simon, 1963) and CYC (Lenat, 1995). Except serving as a physical container of the system, the body of such a system has little to do with the content and behavior of the system. Even in robotics, where the role of body cannot be ignored, the traditional approach works in a Sense-Model-Plan-Act (SMPA) framework, in which the robot acts according to an internal “world model”, a symbolic representation of the world (Nilsson, 1984; Brooks, 1991a).

As a reaction to the problems in this tradition, the ‘embodied’ approach criticizes the traditional approach as being ‘disembodied’, and emphasizes the role of sensorimotor experience to intelligence and cognition. Brooks’ behavior-based robots have no representation of the world or the goal of the system, since “the world is its own best model” (Brooks, 1991a), so the actions of the robot are directly triggered by corresponding sensations. According to Brooks, “In order to really test ideas of intelligence it is important to build complete agents which operate in dynamic environments using real sensors. Internal world models which are complete representations of the external environment, besides being impossible to obtain, are not at all necessary for agents to act in a competent manner.” (Brooks, 1991a)

Therefore, as far as the current discussion is concerned, ‘embodiments’ means the following two requirements:

Working in real world: “Only an embodied intelligent agent is fully validated as one that can deal with the real world” (Brooks, 1991a), since it is more realistic by taking the complex, uncertain, real-time, and dynamic nature of the world into consideration (Brooks, 1991a; Pfeifer and Scheier, 1999).

Having grounded meaning: “Only through a physical grounding can any internal symbolic or other system find a place to bottom out, and give ‘meaning’ to the processing going on within the system” (Brooks, 1991a), which supports content-sensitive processing (Anderson, 2003), and solves the “symbol grounding” problem (Harnad, 1990).

Though this approach has achieved remarkable success in robotics, it still has difficulty in learning skills and handling complicated goals (Anderson, 2003; Brooks, 1991a; Murphy, 2000).
**Embodiment and Robot**

Though the embodiment school has contributed good ideas to AI research, it also has caused some misconceptions.

In the context of AI, it is often suggested, explicitly or implicitly, that only robotic systems are “embodied”, while systems implemented in conventional computer are “disembodied”. This opinion is problematic. As long as a system is implemented in a computer, it has a body — the hardware of the computer. Even though sometimes the system does not have a piece of dedicated hardware, it still stays in a body, the physical devices that carry out the corresponding operations. For instance, a laptop computer obviously has a body, on which all of its software run.

Though the above statement sounds trivially true, some people may reject it by saying that in this context, a “body” means something that have real sensorimotor mechanism, as suggested in (Brooks, 1991a; Brooks, 1991b; Pfeifer and Scheier, 1999). After all, robots have sensors and actuators, while laptop computers do not, right? Though this is indeed how we describe these two types of system in everyday language, this casual distinction does not make a fundamental difference. As long as a system interacts with its environment, it has sensors and actuators, that is, input and output devices. For a laptop computer, its sensors include keyboard and touch-pad, and its actuators include screen and speaker, while the network connection serves as both. These devices are different from the ones of robots in the type, range, and granularity of signals accepted/produced, but they are no less “real” as sensorimotor devices. Similarly, computer input and output operations can be considered as “perception” and “action”, in the broad sense of the words.

How about the claim that only robots interact with the “real” world? Once again, it is a misleading claim, because the environment of other (non-robotic) systems are no less “real” — at least the human users who use the computer via the input/output decides are as real as the floor the robots run on! After all, to a robot, the “world” it can perceive is still limited by the function of its sensorimotor devices.

A related distinction is between “physical agents” (like robots) and “virtual agents” (like chatbots). They are clearly different, but the difference is not that the latter does not run in a physical device or does not interact with its environment via physical processes — the electric currents carrying the input/output signals for a chatbot are as “physical” as the lights going into the visual sensor of a robot.

The above misconceptions usually come from the opinion that though an ordinary computer has a hardware body and does interact with its environment, the interaction is symbolic and abstract, and therefore is fundamentally different from the physical and concrete interaction between a robot and its environment. However, this opinion is an misunderstanding itself.

In the context of the current discussion, there is no such a thing as “purely symbolic and abstract interaction”. Every interaction between every computer and its environment is carried out by some concrete physical process, such as pressure on a key, movement on a touch-pad, light change on a monitor, electronic flow in a cable, and so on. What is ‘symbolic’ and ‘abstract’ is not such a process itself, but the traditional description about it, where the details of the underlying physical process is completely omitted. On this topic, the difference between a computer and a robot is not really in the system themselves, but in the usual ways to treat them.

Now some reader may think that this paper is another defense of the symbolic AI school against the embodiment school, like (Vera and Simon, 1993), since it dismisses the embodiment approach by saying that what it demands are already there all the time. This is not the case. What this paper wants to do is actually to strengthen the embodiment argument, by rejecting certain common misunderstandings and focusing on the genuine issues.

Though every computer system has a body, and does interact with its environment, there is indeed something special about robots: a robot directly interacts with the world without human involvement, while the other systems mainly interact with human users. As argued above, here the difference is not whether the world or the sensor/actuator is “real”. Instead, it is that the human users are tolerant to the system, while the non-human part of world is not. In robotics, “There is no room for cheating” (Brooks, 1991a) — a robot usually has to face various kinds of uncertainty, and to make real-time response. On the contrary, in other AI systems there are various assumptions on what types of input are acceptable, and on how much time-space resources are required for a certain computation, that the users have got used to gratify.

Therefore, the “real world” requirement is really about whether the assumptions on environment are “realistic”, by keeping its complexity, uncertainty, and resource-restriction. Under this interpretation, “be real” is applicable not only to robots, but also to almost all AI systems, since in most realistic situations, the system has insufficient knowledge (various uncertainties) and insufficient resources (time-space restriction), with respect to the problems. It is only that traditional AI systems have the option to “cheat” by only accepting an “idealized” version of a problem, while robotic systems usually do not have such an option.

The traditional symbolic AI systems are indeed disembodied. Though every AI system has a body (with real sensors and actuators) and interacts with the real world, in traditional AI systems these factors are all ignored. Especially, in the internal representation of the world in such a system, the meaning of a symbol is determined by its denotation in the world, and therefore have little to do with the system’s sensorimotor experience, as well as the bias and restriction imposed by the system’s body. For example, if the meaning of symbol “Garfield” is nothing but a cat existing in the world, then whether a system using the symbol can see or touch the cat does not matter. The system does not even need to have a body (even though it does have one) for the symbol to have this meaning. This is not how meaning should be handled in intelligent systems.

Based on the above analysis, the two central requirements of embodiment can be revised as the following:

**Working in real world:** An intelligent system should be designed to handle various types of uncertainty, and to work in real time.
Having grounded meaning: In an intelligent system, the meaning of symbols should be determined by the system’s experience, and be sensitive to the current context.

This version of embodiment is different from the Brooks-Pfeifer version, in that it does not insist on using robots to do AI (though of course it allows that as one possibility). Here “embodiment” no longer means “to give the system a body”, but “to take the body into account”. According to this opinion, as long as a system is implemented, it has a body; as long as it has input/output, it has perception/action. For the current discussion, what matters is not the physical properties of the system’s body and input/output devices, but the experience they provide to the system. Whether a system is “embodied” is determined by whether the system is adaptive to its experience, as well as whether there are unrealistic constraints on its experience.

Many traditional AI systems are disembodied, not because they are not implemented as robots, but because the symbols in them are understood as labels of objects in the world (therefore are experience-independent), and there are strong constraints on what the system can experience. For example, the users should not feed the system inconsistent knowledge, or ask questions beyond its knowledge scope. When these events happen, the system either refuses to work or simply crashes, and the blame falls on the user, since the system is not designed to deal with these situations.

**Embodiment in NARS**

To show the possibility of achieving embodiment (as interpreted above) without using a robot, an AGI project, NARS, is briefly introduced. Limited by the paper length, here only the basic ideas are described, with reference to detailed descriptions in other publications.

NARS is a general-purpose AI system designed according to the theory that “intelligence” means “adapting to environment and working with insufficient knowledge and resources” (Wang, 2006). Since the system is designed in the reasoning system framework, with a formal language and a set of formal inference rules, at the first glance it looks just like a “disembodied” traditional AI system, though this illusion will be removed, hopefully, by the following description and discussion.

At the current stage of development, the interaction between NARS and its environment happens as input or output sentences of the system, expressed in a formal language. A sentence can represent a judgment, a question, or a goal. As input, a judgment provides the system new knowledge to remember, a question requests the system to find an answer according to available knowledge, and a goal demands the system to achieve it by carrying out some operations. As output, a judgment provides an answer to a question or a message to other systems, a question or a goal asks help from other systems in the environment to answer or achieve it. Over a period of time, the stream of input sentences is the system’s experience, and the stream of output sentences is the system’s behavior.

Since NARS assumes insufficient knowledge, there is no constrain on the content of its experience. New knowledge may conflict with previous knowledge, no knowledge is absolutely certain, and questions and goals may be beyond the current knowledge scope. Consequently, the system cannot guarantee the absolute correctness of its conclusions, and its predictions may turn out to be wrong. Instead, the validity of its inference is justified by the principle of adaptation, that is, the conclusion has the highest evidential support (among the alternatives), according to the system’s experience.

Since NARS assumes insufficient resources, the system is open all the time to new input, and processes them in real-time. So the system cannot simply process every problem exhaustively by taking all possibilities into consideration. Also, it has to manage its storage space, by removing some data whenever there is a shortage of space. Consequently, the system cannot guarantee the absolute optimum of its conclusions, and any of them may be revised by new information or further consideration. Instead, the validity of its strategy is also justified by the principle of adaptation, that is, the resources are allocated to various activities to achieve the highest overall efficiency (among the alternatives), according to the system’s experience.

The requirement of embodiment follows from the above assumption and principle. The assumption on the insufficiency in knowledge and resources puts the system in a realistic environment, where it has to deal with various types of uncertainty, and handle tasks in real-time. The system does not have the knowledge and resources to build a model of the world, then to act accordingly. Instead, its knowledge is nothing but summary of its past experience, which guides the system to deal with the present, and be prepared for the future. There is an internal representation in the system, though it is not a representation of the world, but a representation of the experience of the system, after summarization and organization. The symbols in the representation have different meaning to the system, not because they refer to different objects in the world, but because they have played different roles in the system’s experience.

Concretely, the meaning of a concept in NARS is determined by its experienced relation with other concepts. That is to say, what “Garfield” means to (an implementation of) NARS is not decided by an object labeled by that term, but by what the system knows about “Garfield”. Given the resources restriction, each time the concept “Garfield” is used in the system, only part of its relations are taken into consideration. Therefore, what the term means to the system may (more or less) change from time to time, and from situation to situation, though not arbitrarily.

The details of this “experience-grounded semantics” is explained and discussed in (Wang, 2005; Wang, 2006). Though many people have argued for the importance of experience in intelligence and cognition, no other work has explicitly and formally defined the central semantic notions ‘meaning’ and ‘truth-value’ as functions of the system’s experience, and specified the details in their computational implementation.

How about the sensorimotor aspects of the meaning? In a broad sense, all knowledge (directly or indirectly) comes from the system’s experience, which initially comes through sensorimotor devices of the system. If we use the term to re-
fer to non-linguistic experience, then in NARS it is possible to link “Garfield” to related visual images and operation sequences, so as to enrich its meaning. However, it is important to understand that both linguistic experience and non-linguistic experience are special cases of experience, and the latter is not more “real” than the former. In the previous discussions, many people implicitly suppose that linguistic experience is nothing but “Dictionary-Go-Round” (Harnad, 1990) or “Chinese Room” (Searle, 1980), and only non-linguistic sensorimotor experience can give symbols meaning. This is a misconception coming from traditional semantics, which determines meaning by referred object, so that an image of the object seems to be closer to the “real thing” than a verbal description. NARS’ experience in Chinese is different from the content of a Chinese-Chinese dictionary, because a dictionary is static, while the experience of a system extends in time, in which the system gets feedback from its environment as consequences of its actions, i.e., output sentences in Chinese. To the system, its experience contains all the information it can get from the environment. Therefore, the system’s processing is not “purely formal” in the sense that the meaning of the symbols can be assigned arbitrarily by an outside observer. Instead, to the system, the relations among the symbols are what give them meaning. A more detailed discussion on this misconception can be found in (Wang, 2007), and will not be repeated here.

In summary, NARS satisfies the two requirements of embodiment introduced previously:

**Working in real world:** This requirement is satisfied by the assumption of insufficiency in knowledge and resources.

**Having grounded meaning:** This requirement is satisfied by the experience-grounded semantics.

### Difference in Embodiment

Of course, to say an implementation of NARS running in a laptop computer is “already embodied”, it does not mean that it is embodied in exactly the same form as a human mind operating in a human body. However, here the difference is not between “disembodied” and “embodied”, but between different forms of embodiment.

As explained previously, every concrete system interacts with its environment in one or multiple modalities. For a human being, major modalities include vision, audition, tactile, etc.; for a robot, they include some human-like ones, but also non-human modalities like ultrasonic; for an ordinary computer, they directly communicate electronically, and also can have optional modalities like tactile (keyboard and various pointing devices), audition (microphone), vision (camera), though they are not used in the same form as in a human body.

In each modality, the system’s experience is constructed from certain “primes” or “atoms” that is the smallest units the system can recognize and distinguish. The system’s processing of its experience is usually carried out on their compound “patterns” that are much larger in scale, though short in details. If the patterns are further abstracted, they can even become modality-independent “symbols”. This is the usual level of description for linguistic experience, where the original modality of a pattern, with all of its modality-specific details, is ignored in the processing of the message. However, this treatment does not necessarily make the system disembodied, because the symbols still comes from the system’s experience, and can be processed in an experience-dependent manner.

What makes the traditional symbolic AI system disembodied is that the symbols are not only abstracted to become *modality-independent*, but also *experience-independent*, in the sense that the system’s processing of the symbol is fully determined by the system’s *design*, and have little to do with its *history*. In this way, the system’s body becomes completely irrelevant, even though literally speaking the system exists in a body all the time.

On the contrary, linguistic experience does not exclude the body from the picture. For a system that only interact with its environment in a language, its experience is linguistic and amodal, in the sense that the relevant modality is not explicitly marked in the description of the system’s experience. However, what experience the system can get is still partially determined by the modality that carries out the interaction, and therefore, by the body of the system. As far as the system’s behavior is *experience-dependent*, it is also *body-dependent*, or *embodied*.

Different bodies give a system different experiences and behaviors, because they usually have different sensors and operators, as well as different sensitivity and efficiency on different patterns in the experience and the behavior. Consequently, even when they are put into the same environment, they will have different experience, and therefore different thoughts and behaviors. According to experience-grounded semantics, the meaning of a concept depends on the system’s experience on the concept, as well as on the possible operations related to the concept, so any change in the system’s body will more or less change the system’s mind.

For example, at the current stage, the experience of NARS is purely linguistic, so the meaning of a concept like ‘Garfield’ only depends on its experienced relations with other concepts, like ‘cat’, ‘cartoon character’, ‘comic strip’, ‘lazy’, and so on. In the future, if the system’s experience is extended to include visual and tactile components, the meaning of ‘Garfield’ will include additional relations with patterns in those modalities, and therefore become closer to the meaning of ‘Garfield’ in a typical human mind. Therefore, NARS implemented in a laptop and NARS implemented in a robot will probably associate different meaning to the same term, even though these meanings may have overlap.

However, it is wrong to say that the concept of ‘Garfield’ is *meaningful or grounded* if and only if it is used by a robot. There are two common misconceptions on this issue. One is to only take sensorimotor experience as real, and refuse to accept linguistic experience; and the other is to take human experience as the standard to judge the intelligence of other systems. As argued previously, every linguistic experience must be based on some sensorimotor experience, and though the latter is omitted in the description, it does not make the former less ‘real’ in any sense. Though “behave according
to experience” can be argued to be a necessary condition of being intelligent (Wang, 2006), to insist the experience must be equal to or similar to human experience leads to an anthropocentric understanding of intelligence, and will greatly limit our ability to build, and even to imagine, other (non-human) forms of intelligence (Wang, 2008).

In the current AI field, very few research projects aims at accurately duplicating human behaviors, that is, passing the Turing Test. It is not only because of the difficulty of the test, but also because it is not a necessary condition for being intelligent, which was acknowledged by Turing himself (Turing, 1950), though often forgot by people talking about that article. Even so, many outside people still taking “passing the Turing Test” as the ultimate goal, or even the definition, of AI. This is why the proponents of the embodied view of human cognition often have negative view on the possibility of AI (Barsalou, 1999; Lakoff and Johnson, 1998). After identifying the fundamental impacts of human sensorimotor experience on human concepts, they see this as counter evidence for a computer to form the same concepts, without a human body. Though this conclusion is correct, it does not mean AI is impossible, unless “artificial intelligence” is interpreted as “artificial human intelligence”, that is, the system not only follows the general principles associated with ‘intelligence’, but also have the same concepts as a normal human being.

Because of the fundamental difference between human experience and the experience an AI system can have, the meaning of a word like ‘Garfield’ may never be the same in these two types of system. If AI aims at an accurate duplication of the contents of human categories, then we may never get there, but if it only aims at relating the contents of categories and the experience of the system in the same way as in the human mind, then it is quite possible, and that is what NARS attempts to achieve, among other things.

When people use the same concept with different meanings, it is usually due to their different experience, rather than their different intelligence. If this is the case, then how can we expect AI systems to agree with us on the meaning of a word (such as “meaning”, or “intelligence”), when we cannot agree on it among ourselves? We cannot deny the intelligence of a computer system just because it uses some of our words in a way that is not exactly like human beings.

Of course, for many practical reasons, it is highly desired for the concepts in an AI system to have similar meaning as in a typical human mind. In those situations, it becomes necessary to simulate human experience, both linguistic and non-linguistic. For the latter, we can use robots with human-like sensors and actuators, or simulated agents in virtual worlds (Bringsjord et al., 2008; Goertzel et al., 2008). However, we should understand that in principle, we can build fully intelligent systems, which, when given experience that is very different from human experience, may use some human words in non-human ways. After all, “to ground symbols in experience” does not mean “to ground symbols in human experience”. The former is required for being intelligent, while the latter is optional for being intelligent, though maybe desired for certain practical purposes.

**Conclusion**

Embodiment is the request for a system to be designed to work in a realistic environment, where its knowledge, categories, and behavior all depend on its experience, and therefore can be analyzed by considering the interaction between the system’s body and the environment.

The traditional symbolic AI systems are disembodied, mainly because of their unrealistic assumptions about the environment, and their experience-independent treatment of symbols, categories, and knowledge.

Though robotic research makes great contribution to AI, being a robot is neither a sufficient nor a necessary condition for embodiment. When proposed as a requirement for all AI systems, the requirement of embodiment should not be interpreted as “to give the system a body”, or “to give the system a human-like body”, but as “to make the system behave according to its experience”. Here “experience” includes linguistic experience, as a high-level description of certain underlying sensorimotor activity.

The practice in NARS shows that embodiment can be achieved by a system where realistic assumption about the environment is made, such as “the system has insufficient knowledge/resources with respect to the problems the environment raises”, and the symbols in the system can get their meaning from the experience of the system, by using an experience-grounded semantics.

Though a laptop computer always has a body, a system running in this laptop can be either “embodied”, or “disembodied”, depending on whether the system behaves according to its experience.

Different bodies give systems different possible experiences and behaviors, which in turn lead to different knowledge and categories. However, here the difference is not between intelligent systems and non-intelligent ones, but among different types of intelligent systems.

Given the fundamental difference in hardware and experience, we should not expect AI systems to have human concepts and behaviors, but the same relationship between their experience and behavior, that is, being adaptive, and working with insufficient knowledge and resources.

**Acknowledgment**

The author benefits from a related discussion with Ben Goertzel and some others on the AGI mailing list, as well as from the comments of the anonymous reviewers.

**References**

Anderson, M. L. (2003). Embodied cognition: A field guide. Artificial Intelligence, 149(1):91–130.

Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22:577–609.

Bringsjord, S., Shilliday, A., Taylor, J., Werner, D., Clark, M., Charpentie, E., and Bringsjord, A. (2008). Toward logic-based cognitively robust synthetic characters in digital environments. In Artificial General Intelligence 2008, pages 87–98, Amsterdam. IOS Press.
Brooks, R. A. (1991a). Intelligence without reason. In Proceedings of the 12th International Joint Conference on Artificial Intelligence, pages 569–595, San Mateo, CA. Morgan Kaufmann.

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

Goertzel, B., Pennachin, C., Geissweiller, N., Looks, M., Senna, A., Silva, W., Heljakka, A., and Lopes, C. (2008). An integrative methodology for teaching embodied non-linguistic agents, applied to virtual animals in Second Life. In Artificial General Intelligence 2008, pages 161–175, Amsterdam. IOS Press.

Harnad, S. (1990). The symbol grounding problem. Physica D, 42:335–346.

Lakoff, G. and Johnson, M. (1998). Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. Basic Books, New York.

Lenat, D. B. (1995). Cyc: A large-scale investment in knowledge infrastructure. Communications of the ACM, 38(11):33–38.

Murphy, R. R. (2000). An Introduction to AI Robotics (Intelligent Robotics and Autonomous Agents). MIT Press, Cambridge, Massachusetts.

Newell, A. and Simon, H. A. (1963). GPS, a program that simulates human thought. In Feigenbaum, E. A. and Feldman, J., editors, Computers and Thought, pages 279–293. McGraw-Hill, New York.

Newell, A. and Simon, H. A. (1976). Computer science as empirical inquiry: symbols and search. Communications of the ACM, 19(3):113–126.

Nilsson, N. J. (1984). Shakey the robot. Technical Report 323, SRI AI Center, Menlo Park, CA.

Pfeifer, R. and Scheier, C. (1999). Understanding intelligence. MIT Press, Cambridge, Massachusetts.

Searle, J. (1980). Minds, brains, and programs. The Behavioral and Brain Sciences, 3:417–424.

Turing, A. M. (1950). Computing machinery and intelligence. Mind, LIX:433–460.

Vera, A. H. and Simon, H. A. (1993). Situated action: A symbolic interpretation. Cognitive Science, 17(1):7–48.

Wang, P. (2005). Experience-grounded semantics: a theory for intelligent systems. Cognitive Systems Research, 6(4):282–302.

Wang, P. (2006). Rigid Flexibility: The Logic of Intelligence. Springer, Dordrecht.

Wang, P. (2007). Three fundamental misconceptions of artificial intelligence. Journal of Experimental & Theoretical Artificial Intelligence, 19(3):249–268.

Wang, P. (2008). What do you mean by “AI”? In Artificial General Intelligence 2008, pages 362–373, Amsterdam. IOS Press.