The World According to a Humanoid Robot

Erich Prem*
The Austrian Research Institute for Artificial Intelligence
Schottengasse 3, A-1010 Wien, Austria
erich@ai.univie.ac.at

Abstract

The new field of embodied Artificial Intelligence deals with the construction of robotic systems that exhibit high interaction dynamics with the world and intelligent behavior. The construction of humanoid robots is particularly challenging and poses many foundational and methodological questions. Two such questions are "What does the world look like for a human-like artificial systems?" and "Why should we care?". It is argued that the study of autonomous intelligent systems must be based on a new understanding of the relation between the cognizer and its environment. We describe building blocks of such an approach to the construction of humanoid systems.

1 Introduction

1.1 Artificial Intelligence

On the face of it, the construction of a humanoid robot is a completely hopeless endeavour today. We are neither in possession of a clear theory about how the mind works, nor are we currently equipped with a technology to realize all the "computational" functions of the human brain. Additionally, there are hardly any artificial mechanical devices that can realize the skill of human body-parts paired with a comparable elegance in design. Nevertheless, there are people striving for the goal of building an artificial system that exhibits at least some of the physical and mental properties of human beings. The reason for this stems not only from an interest in technological challenges, but also from a desire to discover how the human mind works.

*Presented at robots@aec.at, Workshop on Teleoperation and Robotics, June, Linz, Austria, 1997. To appear in "Schriftenreihe der OCG", Austrian Computer Society, 1997.

One scientific discipline in which both the technical and the natural scientific challenge always existed is the field of Artificial Intelligence. For more than thirty years now AI researchers have built computer programs that can solve some tasks which, when solved by humans, are believed to require intelligence. However, bodily behavior, i.e. the construction of working robots, remained a very difficult and hardly solved problem. Some scientists nowadays believe that the reason for this lies not only in the lack of human-like arms or legs, but in a wrong construction of the fundamental theories underlying traditional Artificial Intelligence.

These classic approaches centered around computational readings of what it means to be intelligent. Consequently, many formal and mathematical theories were developed to make computers smart. The general approach often consisted in analyzing the input to a system so as to construct complex internal representations that could later be used for the calculation of the desired output. This also used to be the general approach to the construction of robots. Typically, input from many different sensors was evaluated by a complex analysis and resulted in some kind of internal model of the world. This model was then used by a planning algorithm to calculate and later execute the proper system actions (cf. [19]). This approach is depicted in the left part of Fig. 1 (a) [1, 2].

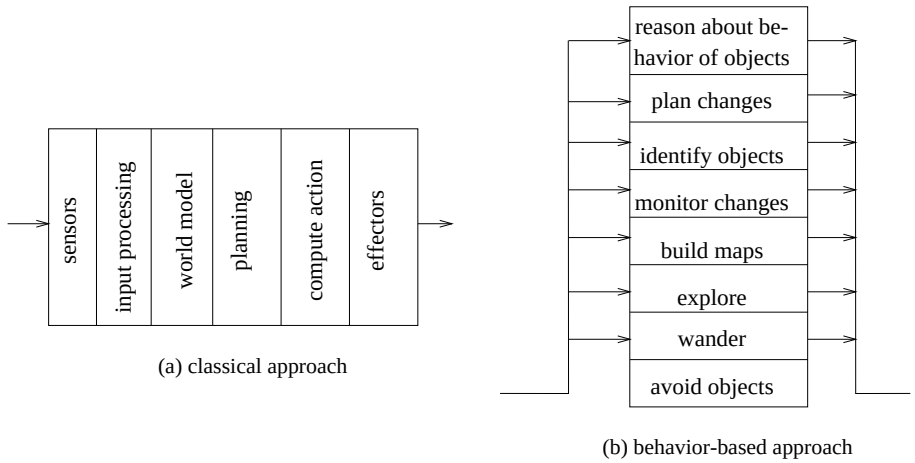

Figure 1: Classical (left) and behavior-based robot architecture.

Nowadays it is well-known that such an approach to the construction of robots comes with many problems. The functional separation of sensor interpretation, world model construction, planning and execution introduces many time delays and reduces the system dynamics, i.e. the *speed* with which

it can interact with its environment. This, however, is not what one would expect from a humanoid robot. Ordinary people attribute intelligence also based on the speed with which a certain action is taken or a correct reply to a question is given. This discrepancy recently lead to the development of the field of embodied Artificial Intelligence.

1.2 Embodied Artificial Intelligence

Embodied AI deals with the construction of robotic systems that exhibit high interaction dynamics with their environment. Proponents of embodied AI believe that "intelligence is determined by the dynamics of the interaction with the world" [2]. Besides of the aspect of mere interaction speed there is another important implication of this believe. The system behavior is not a purely compuational task that happens to move a body around. Quite to the contrary, the physical interaction plays an important role in the realization of the desired behavior. To the proponent of embodied AI it is unimportant whether the robot shows intelligent behavior due to a computational procedure or due to its clever physical design. Moreover, real physical embodiment (as opposed to a mere simulation) ensures a certain realism when it comes to testing a newly devised algorithm [15].

In the last decade, many robots have been developed within this framework (sometimes called "behavior-based" robotics). Their functionality ranges from grass-mowing, vacuum-cleaning, to soda-can collection [3]. The robots autonomously perform their tasks at relatively high interaction dynamics in unstructured environments. The key to the realization of this new kind of robots was to abandon the idea of complex sensor evaluation and a central world model. Instead, these robots are equipped with relatively simple sensors, e.g. bumpers, infrared or subsonic range finders, and a few specialized sensors that are custom-made for the task at hand. As an example one may consider the soda-can detection mechanism of one of these robots [5]. The meaning of this laser-range finder is physically restricted so that it can be interpreted as "there is a soda-can in front of the gripper", at least in typical office environments. The physical constraint allows a quick interpretation of the sensor and the introduction of reactive behaviors that control the robot.

The new system architecture of behavior-based robotics is depicted in Fig. 1 b. There is no centralized model of the world in this architecture. Instead, single layers of basic behaviors receive inputs from the environment. All theses behaviors run in parallel and have the possibility to create commands that go to the effectors. Behaviors from higher levels may suppress

lower-level behaviors. In many existing architectures, the input to each level stems from a specific sensor. For example, the input to "avoid objects" could stem from a simple IR range finder.

1.3 Biomimetic robotics

The methodology for the construction of this kind of behavior-based robots can be more easily understood within a biological or ethological framework. The theoretical biologist Jakob von Uexküll has described the "action circuit" as the fundamental building block of animal behavior [20]. It is depicted in Fig. 2. Uexküll argued that the explanation of animal behavior

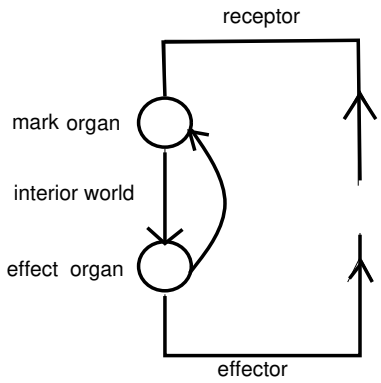

Figure 2: Action circuit as described by von Uexküll.

must be based on an understanding of how the things around the animal are embedded in an action and how this action is embedded in a purposeful interaction with the world. From Fig. 2 it can be seen that the interior world of the animal is, of course, dependent on the structure of the sensory organs. But since the animal acts upon its environment, the internal world is equally dependent on the animal's actions. It is the interaction with its environment that forms the internal world of the animal.

The formation of sensory experience is not only based on the system's interaction. Even more importantly, the interaction has a specific purpose. Such a purpose turns the objects encountered and acted on by the system from a collection of merely causally operating parts of physical entities into a meaningful assembly of things which are integrated in a purposeful whole. The essential point is to understand how the thing is embedded in an action and how this action is embedded in a purposeful interaction with the world. In order to fully understand the system's world, our task consists in the

dissection of the functional world (i.e. the whole of the subject's functional circuits).

Take as an example the behavior of a tick. From experiments we know that the tick bites everything which has a superficial temperature of 37° C and emits butyric acid. That means, within the context of feeding, it is only chemical concentration an temperature that guide the animal's actions. Within the feeding context, the world according to the tick is only warm or cold and has or has not a specific "smell".

Coming back now to our robotic example the behavior-based methodology consists in describing a set of functional interactions of the robot and its environment. While these functional circuits could be "fleeing, feeding, mating" in the case of an animal, they are often "homing, wandering, collecting" in the case of a robot. From these functional descriptions it follows which kind of physically restricted sensors must be constructed to achieve the desired functionality. But note that this also means to design *the world according to the robot.*

2 Studying cognition with "Cog"

2.1 The humanoid robot "Cog"

Many philosophers (e.g. [21]) have argued that the human body is of central importance to human thinking, much more than traditional AI has usually accepted. It is well-known that we use many body-based metaphors in our language and it is suspected that these metaphors also play an important role in abstract reasoning [11]. If the possession of a physical body is in any sense constitutive for intelligent behavior, than it follows that in order to study intelligence, a system is needed that has a real body [16]. One system that is currently under development at the MIT AI laboratory is the humanoid robot "Cog" [4]. The purpose of "Cog" is not only the construction of a working system, but also to gain insights into the nature of processes that underlie human cognition.

The robot is mounted to an immovable platform. It is supposed to resemble a human from the waist up, torso and head each having three degrees of freedom. Currently, one arm (6 degrees of freedom) is mounted to the robot's body. A special human-like actuator system used for moving the arm consists of spring-coupled (series elastic) actuators [22]. The hand consists of a 4-finger manipulator with four motors, 36 exteroceptor and proprioceptor sensors [14] controlled by an on-palm micro-controller. Cog's head carries four black and white cameras which together form an active

vision system, each eye having two degrees of freedom. The system can be best described as consisting of two pairs of cameras, one pair forming one "eye". Cog is also equipped with an auditory perception system [10]. The control system for Cog consists of a large number of micro-controller boards [12].

The robot is capable of tracking moving objects with its artificial eyes at a speed comparable to humans. It has learned to map vision coordinates on spatial coordinates and can use this information for the control of its motors. It can grasp for objects, learn to distinguish them based on their elasticity and give them back to people who play with the robot. The interaction of the robot with its environment is currently triggered by external events. For example, the robot will automatically turn its eyes (and consequently its head and neck) towards any motion within the vision field. It should be noted that although the system is very reactive (as opposed to cognitive) it has a rather natural appearance during interaction with people.

2.2 Implications for the study of cognition

The existence of a physical robot like Cog allows the scientists to study the dynamics of a human-like body. The problem of human intelligence, i.e. cognition, can now be investigated as the problem of generating physical behavior. As a consequence, intelligence is not viewed in separation of its bodily substrate. Cognition is a timely process driven by forces internal and external to the system. It happens in close interaction with the world, often in order to manipulate the world [16].

In this approach to Cognitive Science the modeling of cognitive phenomena starts with an analysis of bodily processes as well as mere "mental" ones. The result is a list of model requirements which contains physical and time constraints. The properties of these models are very different from models in traditional AI, where dynamical aspects are usually only regarded from a mainly computational point of view. However, real body dynamics can be very different from simulated ones. There may be no in-principle difference between two simulations running at different speeds. There will, however, often be major physical differences between a slow and a fast robot. This difference can also imply severe consequences for the kind of control mechanisms that can make the system work.

In Cog, the control systems are of a highly parallel architecture. The control system is similar to the one depicted in Fig. 1 with different behaviors running on separate processors. The overall behavior of Cog is therefore very unpredictable. It reacts to different stimuli in its environment, but it can

not always be told what exactly generated a specific form of behavior. The reason for this lies in the fact that the different parts of the body physically interact. This can sometimes make the reproduction of an observed behavior nearly impossible.

Critics of this approach to the study of cognition argue that such a system may be at the edge of technology, but that it has hardly anything to do with intelligence. After all, event-triggered reactive behavior is not what chess players or highly intelligent mathematicians do. In what follows, I will give arguments why the embodied approach to intelligence does justice to human behavior and what this implies for the research methodology of AI.

3 Being in the world

A proper understanding of the relation between an intelligent cognizer and its environment is central to the embodied AI approach to cognitive science. Studying this relation means to make efforts so as to understand "the world according to the humanoid robot". In the example of the tick behavior, we have seen that there are only two important features of the world within the feeding context (temperature and chemical concentration). When constructing a robot, we must, of course, start with the functions and from this derive the interaction circuits of the robot and design the sensors and behaviors accordingly.

Such a methodology must therefore start with a study of what there is in the robot's environment and how it is perceived and acted upon by the system. This leads to a new system ontology, i.e. to a reformulation of what there is in the world in terms of the autonomous agent. This first task consists in the description of the world in terms that can be made operational through the robot's sensors and actuators.

As innocent as this descriptional framework may look, it has a rather strong influence on the system's representational framework. It is now likely that sensory impressions of the system are categorized in classes that form items of the same usefulness to the system. This is a direct consequence of the "action circuit" that defines things by means of their sensory properties dependent on the outcome of interaction. Since this interaction serves some specific purpose for the system, the "things" will become mainly defined by means of their functions for the system.

In the same sense that "chairs" are properly described by their function "for sitting" for humans, objects in an embodied system's environment will now be classified due to their functional properties. It is clear that such

a representational frame can be conceptually opaque in relation to human concepts.

In his interpretation of Merleau-Ponty's philosophy, Dreyfus points out that there are at least two important ways in which the human mind reflects its physical embodiment ([8] and also [13]). A part of his argument can be easily translated to robotics. (i) Firstly, the actual shape and capacities of the robot determine the adequateness of objects around us for certain types of interaction. Chairs are for sitting, while coffee cups are not. (ii) Secondly, the cultural world is also correlative with our body. Chairs do have a certain culturally dependent meaning in our Western world, while they may be regarded useless by some primitive race. (iii) Thirdly, bodily functions determine the kind of purposeful interactions with the world that underlie so many actions. In this context, chairs provide recreation to the exhausted.

The translation of this line of reasoning to robotics looks as follows: (i) The first of these three arguments is particularly important. It implies the actual necessity of a system to resemble the shape of the human body. Without this shape, the world according to the human will always be different from the world according to the robot. (ii) The second argument is already more problematic, because it implies that the robot must "grow up" in a culturally biased environment and be able to adapt to it. However, this is another reason for the human form. Natural interaction with the robot seems only possible with something resembling a human. (iii) The last of the three arguments, however, is hard to realize for any electro-mechanical system that does not need food or drink, pleasure or rest–at least not in any conventional human sense. This implies that such a system's artificial needs and drives (e.g. "low battery") will always remain artificial metaphors for their natural counterparts (e.g. "hunger").

3.1 Robotic versus human action

Still, there remains the doubt whether the kind of reactive interaction of "Cog" is an adequate way to study human behavior. I would like to argue that it indeed is and that the approach could correct a long existing scientific misinterpretation of everyday human activity.

The world according to the humanoid robot Cog consists of only a few meaningful things, for example, moving objects, things in its hand, etc. The way in which Cog deals with these items is not based on any further in-depth analysis of sensory properties or of functions that these things may fulfill. Instead, Cog deals with these items in an immediate, event-triggered way.

For Cog, they are core elements of rather simple action circuits. *Dealing* with these things is the primary way of encountering them, not some bare sensory perception. Cog thus "floats" from one action to the other, often encounters things in parallel and seems to be continuously coping with the world.

This kind of phenomenological description of Cog's behavior is reminiscent of accounts that the philosopher Martin Heidegger gives for human Beings [9]. His analysis of people's everyday conduct does not start with a prima-facie analysis of how people analyze their environment and look for objects that might fulfill specific purposes. Instead, Heidegger's phenomenological description is more in line with our personal everyday "experience", i.e. that we are usually engaged in everyday activities without ever consciously thinking about what we are doing. Dreyfus, an early critic of traditional AI, based his critique of traditional AI on Heidegger's philosophy [6, 7].

The things in the world according to Cog can be regarded as "equipment" in Heidegger's terminology [17]. This is the kind of objects that we come across in our everyday activity of writing, sewing, working, etc. Entities like pens, needles or hammers are mostly not encountered in some intellectual, detached way. Often, we feel that we are just dealing with them, as if they would lend themselves nicely to the fulfillment of different tasks. In our everyday activity we do not have to think about how to use these things, they are simply "ready-to-hand". Heidegger's famous example is the use of a hammer. When hammering, we do not think about the hammer as a tool that is conventionally used for nailing. Instead, it seems more like the hammer would refer to its purpose of hammering and, moreover, to many other things like nails, wood, etc. that we might find useful when hammering.

> "The world presents itself in the equipmental nexus, in the reference to a previously seen whole." [9, p.75]

It is exactly this kind of purpose-based ontology that is supported by the methodological approach that lies behind the construction of a system like Cog or the more simple behavior-based autonomous systems. The things in the world according to Cog are not just equipped with some predicate that defines what they can be used for. Rather, they *are* things because they are elements in a circuit of purposeful interaction with the world. The fact that the interaction itself is reactive rather than deliberate does not seem to play an important role in this context.

3.2 Detached staring

There is, of course, a sense in which Cog's behavior and attitude towards its environment seems particularly un-humanlike. People do not *always* cope with the world in such a "flowing", continuous and unproblematic action. For example, it may occur that the hammer does not work properly, its head may need to be fixed. This is a situation in which we may be startled for a short period and then quickly find a way to fix the situation. If the problem cannot be solved immediately, however, we actually interrupt our current activity. We may then be forced to look at the hammer in more detail, to investigate its context-free features and properties. This then would be the situation that was addressed and described by traditional AI. In such a situation of breakdown, the intelligent cognizer interacts with its environment in a rather detached, analytic way. It looks for properties and features of the things at hand, indeed the things now turn into property-bearers and become equipped with "functions".

This analytic attitude towards the objects in the world also lies at the core of what we mean by "intelligent". The person who can skillfully deal with problematic situations and malfunctioning tools is likely to be called intelligent. However, this kind of dealing with the world is based on the everyday unproblematic coping with our environment. And the logical way to study the latter form of "detached cognition" is to found it on the more basic way of "everyday coping".

3.3 De- and re-contextualizing

Another problem associated with such a phenomenological approach to AI is the fact that the things in the world of the autonomous systems arise from their embedding in a concrete functional circuit and from the system's actions. This implies that the things are always regarded from within a rather specific context. For example, the "ball" for Cog is defined (or comes into existence) based on the functional circuit of playing with people. The ball is offered by a human player, Cog takes the ball and gives it back. This immediate contextual embedding is fundamental for the ball's possibility to make reference to other elements of playful interaction, just as the hammer in Heidegger's example refers to other tools or nails.

It must be pointed out that such an object could come to play completely different roles depending on the kind of situational context. What may be a ball in the playing context could be regarded as a weapon in the context of a defensive action. The question arises, how it then becomes possible that

humans also can take the stance of detached observers of their environment. Or in other terms, how do all all the different aspects of things become integrated into one single object?

We have already seen that in situations of breakdown, there arises a need to look at the ball-thing as a mere de-contextualized object. In such a situation the thing seems to be analyzed regardless of its immediate function, only with reference to its perceptual, physical, or functional properties. But this is not the only situation in which things are looked primarily at from an "objective" point of view.

Another such situation of de-contextualization is communication. When I utter the wish "Give me that ball over there!" it is necessary to refer to something that the other person can understand. This kind of linguistic reference to the "ball" is, of course, not completely deprived of the embedding context, but it must also be based on a common identification of the intended object. An easy approach to such a common understanding is to take the thing's external properties, its perceptual features, because I can hope that they are common to both myself and the other person.

This communicative reference to interpersonally experienced features of things could form the basis for our ability to de-contextualize things into objects. In fact, the communicative aspect is sometimes so strong for us that it lead researchers in AI to neglect situations in which decontexutalized properties of objects disappear, such as in everyday coping. Although embodied AI has not found a way to integrate the different contextual aspects of objects, the "things", into a context-free object, the general methodological approach certainly points to this direction.

4 Conclusion

In this paper I tried to give an account of what the world looks like according to a humanoid robot. The current approach to embodied AI is based on an ethological understanding of an autonomous system's interaction with its environment. In such a view the autonomous system's sketch of the world is based on its purposeful interaction with the world. From an extreme viewpoint one can say that everything there is around the robot stems from an action circuit that plays a role in the context of functional interaction with the world.

This view is surprisingly near to Heidegger's phenomenological account of human beings and addresses one of the most fundamental criticisms of AI research. It enables us to no longer regard human beings as purely intellec-

tual analyzers of their surroundings, but to come closer to the phenomenon of everyday human coping with the world. The price for this is the need to explain why humans are also capable of more detached ways of dealing with the world.

It is easy to see that embodied AI as well as the embodied approach to Cognitive Science both still have a very long way to go. On the other hand, it is clear that regarding cognition as a bodily phenomenon changes the character of our approaches to the study of cognition as well as the phenomena of interest, i.e. other kind of questions are asked. Such a new approach to the study of the human mind can also contribute to radically changed view of the way we think about ourselves. It may scratch at the purely rational hull that Greek philosophers or enlightened scientists have put around the modern human and re-acknowledge some of its non-rational, body-based properties. This, of course, is driven by the desire to better understand its rational and non-rational thinking.

Acknowledgements

The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry for Science and Transport. The research described in this paper has been made possible through a travel fund by the Vienna University of Technology (Hrabak scholarship) and kind support from Rodney A. Brooks, MIT AI Lab.

References

[1] Brooks R.A.: A Robust Layered Control System for a Mobile Robot. AI-Laboratory, Massachusetts Institute of Technology, Cambridge, MA, AI-Memo 864, 1985.

[2] Brooks R.A.: Intelligence Without Representation. In Special Volume: Foundations of Artificial Intelligence, Artificial Intelligence, 47 (1–3), pp. 139–159, 1991.

[3] Brooks R.A.: Intelligence Without Reason. In Proceedings of the 12th International Joint Conference on Artificial Intelligence, Morgan Kaufman, San Mateo, CA pp. 569-595, 1991.

[4] Brooks R.A., Stein L.A.: Building Brains for Bodies. In Autonomous Robots, 1, 7–25, 1994.

[5] Connell J.H.: Minimalist Mobile Robotics, Academic Press, San Diego, 1990.

[6] Dreyfus H.L.: What Computers Can't Do. A Critique of Artificial Reason, Harper & Row, New York, 1972.

[7] Dreyfus H.L.: Being-in-the-world. MIT Press, Cambridge, MA, 1991.

[8] Dreyfus H.L.: The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment, Electronic Journal of Analytic Philosophy, 4 (Spring), 1996.

[9] Heidegger M.: Sein und Zeit. (Being and Time.) Tübingen: Nürnberg, 1927.

[10] Irie R.E., Robust Sound Localization: An Application of an Auditory Perception System for a Humanoid Robot, S.M. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, 1995.

[11] Johnson M.: The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. University of Chicago Press, Chicago, 1987.

[12] Kapogiannis, E.: Design of a large scale MIMD computer. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, 1994.

[13] Loren L.A., Dietrich E.: Merleau-Ponty, Embodied Cognition and the Problem of Intentionality, in [18], pp. 345-358, 1997.

[14] Matsuoka, Yoky. Embodiment and Manipulation Learning Process for a Humanoid Hand, S.M. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, 1995.

[15] Prem E.: Grounding and the Entailment Structure in Robots and Artificial Life. In: F. Moràn et al., Advances in Artificial Life, Springer, Berlin, pp. 39–51, 1995.

[16] Prem E.: The implications of embodiment for cognitive theories, Oesterreichisches Forschungsinstitut fuer Artificial Intelligence, Wien, TR-97-11, 1997.

[17] Prem E.: Epistemic Autonomy in Models of Living Systems, Proc. of the Fourth Europ. Conf. on Artif. Life, Brighton, MIT Press/Bradford Books, 1997.

[18] Prem E.(ed.): Special Issue on Epistemological Aspects of Embodied Artificial Intelligence, Cybernetics & Systems, 28, 1997.

[19] Simon H.A.: The Sciences of the Artificial. MIT Press, Cambridge, MA, 1981.

[20] von Uexküll J.: Theoretische Biologie. Suhrkamp, Frankfurt am Main, (1979), 1923.

[21] Varela F.J., Thompson E., Rosch E.: The Embodied Mind, MIT Press, Cambridge, MA, 1991.

[22] Williamson, Matthew. Series Elastic Actuators, MIT AI Lab Tech Report, 1995.