Biologically Inspired Cognitive Architectures (2012) 1, 73–81

Available at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/bica

INVITED ARTICLE

Design principles for biologically inspired cognitive robotics

Jeffrey L. Krichmar

Department of Cognitive Sciences, 2328 Social and Behavioral Sciences Gateway, University of California, Irvine 92697-5100, USA
Department of Computer Science, 2328 Social and Behavioral Sciences Gateway, University of California, Irvine 92697-5100, USA

Received 5 April 2012; received in revised form 28 April 2012; accepted 28 April 2012

KEYWORDS
Cognition;
Computational neuroscience;
Embodiment;
Neurorobots

Abstract
The goals of cognitive robotics are to better understand cognition through the construction of physical artifacts, and to create practical systems that demonstrate cognitive capabilities. I believe for cognitive robotics to move forward, a balanced approach that emphasizes the interaction of brain, body, and environment is necessary. In general, cognitive robots and cognitive architectures focus too much on brain control, and overlook the contributions of morphology to intelligent behavior. On the other hand, the behavior based robotics approach is unbalanced in the opposite direction. For cognitive robotics to move forward, these disparate research communities need to come into balance. The materials, morphology, sensors, actuators, and the nervous system should be balanced and coordinated in their action. In their book, ''How the body shapes the way we think: A new view of intelligence'' (MIT Press, 2007), Pfeifer and Bongard have suggested that intelligent agents should follow a set of design principles that highlight the importance of embodiment and physical interaction with the environment. In the present paper, I apply each of these principles to biologically inspired cognitive robotics and suggest how the field can shift toward better cognitive architectures by adherence to these principles.

Introduction

In the field of cognitive robotics, physical artifacts are constructed to demonstrate a level of cognition. The term cognitive is difficult to define because it has different meaning to different people. In this paper, I will focus on comparing agent behavior to that of biological organisms carrying out cognitive tasks, such as standard tasks of attention,

Address: Department of Cognitive Sciences, 2328 Social & Behavioral Sciences Gateway, University of California, Irvine, CA 92697-5100, USA. Tel.: +1 949 824 5888.
E-mail address: jkrichma@uci.edu

2212-683X/$ - see front matter
http://dx.doi.org/10.1016/j.bica.2012.04.003

decision-making, learning and memory. This allows artificial agents to be compared to a working model. Therefore, throughout the paper the term ''cognition'' could be replaced by ''biologically inspired cognition''. Also, throughout the paper I will use the term ''nervous system'' to describe the system that guides the agent behavior. Although many of the examples given are based on neurobiologically inspired systems, this does not mean to imply that a neurobiologically inspired architecture is the only method for constructing a cognitive robot. But, the neurobiologically inspired approach does allow the mechanism to be compared with empirical data. Despite these restrictions, the arguments below may be applied to cognitive robotics in general.

The field of cognitive robots goes by many names: brain-based devices, cognitive robots, neurorobots, neuromorphic robots, etc. The common goal is twofold: Firstly, developing a system that demonstrates some level of cognitive ability can lead to a better understanding of cognition in general. This idea has been dubbed ''synthetic methodology'' or ''synthetic neural modeling'' and the notion goes back to Grey Walter's Turtles, Braitenberg's vehicles and the Darwin series of automata (Braitenberg, 1986; Edelman et al., 1992; Grey Walter, 1953). The synthetic method has even older roots in psychology and cybernetics (Cordeschi, 2002; Craik, 1967). In general, understanding through building is the goal. Secondly, building a robot or artifact that follows a cognitive model could lead to a system that demonstrates capabilities commonly found in the animal kingdom, but rarely found in artificial systems. While this second goal may have important implications for practical applications, cognitive robotics have not been as successful in achieving this goal as originally hoped. One of the aims of this paper is to discuss some of the reasons why this is the case.

In their book, ''How the body shapes the way we think: A new view of intelligence'', Pfeifer and Bongard put forth an embodied approach to cognition (Pfeifer & Bongard, 2007). Because of this position, many of the robots that they have designed demonstrate ''intelligent'' behavior with limited or non-existent neural processing (Bongard, Zykov, & Lipson, 2006; Iida & Pfeifer, 2004).

In many ways, the field of cognitive robotics and my own work on brain-based devices and neurorobotics might be regarded as the antithesis of Pfiefer and Bongard's position. Our designs are heavy on top-down control and neural processing and light on interaction with the environment. I will discuss why this is the case in detail below. Although they may underemphasize them, cognitive robots do adhere to many of Pfeifer and Bongard's principles. It is my belief that cognitive robots and architectures should attempt to follow these principles.

In the remainder of the paper, I will discuss how each of Pfeifer and Bongard's principles for designing intelligent agents can apply to brain-based or neuromorphic robots. Many readers may not agree Pfeifer and Bongard's point of view. For example, Clark and Grush take the position that cognitive phenomena involves offline-reasoning, vicarious environmental exploration and an internal representation (Clark & Grush, 1999). However, I believe there is value in examining these principles, which arise from behavior based robotics (Arkin, 1998; Brooks, 1991), and seeing how they might be implied to biologically inspired cognitive robotics. I suggest that we in the Biologically Inspired Cognitive Architectures (BICA) community should embrace many of these design principles when developing our systems. I will apologize in advance for not referencing many pertinent robots and systems in the present paper. In general, I will focus on systems that I know at a deep level because either I have had a hand in their design or because I have seen a particular robot in person and discussed the robot with its designers.

Design principles for cognitive robots

Pfeifer and Bongard introduced eight design principles for intelligent agents (Pfeifer & Bongard, 2007). I will follow each of these in turn, and discuss them in light of biologically inspired cognitive robot design.

Agent design principle 1: the three-constituents principle

An intelligent agent should have (1) a defined ecological niche, (2) a defined behavior or task, and (3) an agent design.

The first two constituents define the agent's behavioral task. For example, the niche and behavior of many of our robots has been a controlled laboratory setting. However, much of cognitive science and behavioral neuroscience is conducted in experiments where humans and other animals behave under controlled settings in darkened rooms. Therefore, cognitive robots, which are built to test theories of biological cognition, are often tested in conditions that mimic these experimental paradigms. For example, cognitive robots have replicated standard experimental paradigms such as operant conditioning, fear conditioning, and skill acquisition to better understand learning and memory (Krichmar & Edelman, 2002; McKinstry, Seth, Edelman, & Krichmar, 2008). We have tested our robots in the Morris water maze or the Plus maze to understand how different neural areas contribute to different types of memory (Fleischer, Gally, Edelman, & Krichmar, 2007; Krichmar, Nitz, Gally, & Edelman, 2005).

The main reason for this is to compare the cognitive robot's behavior with biological cognition. This can't be stressed enough. If we are to claim that our robots are cognitive, then we need to test them under conditions by which cognitive scientists test their subjects. However, the flip side of this is that it does not make for exciting robot behavior and it can potentially reveal limitations in the cognitive robot approach. The community needs to demonstrate that cognitive robots can transition to the real world just as humans or animals do when they are outside a laboratory setting.

Some roboticists have been able to demonstrate that their neuromorphic systems are effective outside of laboratory settings. For example, the RatSLAM project has developed accurate cognitive maps of offices and cities based on a hippocampal inspired architecture (Wyeth & Milford, 2009), and our group has shown that a brain-based device can compete on the soccer field (Fleischer et al., 2006).

The third constituent, agent design, is dealt with in cognitive or neuromorphic robots by necessity. If one is testing visual object recognition, then the robot will typically have a vision system, or if one is testing somatosensory

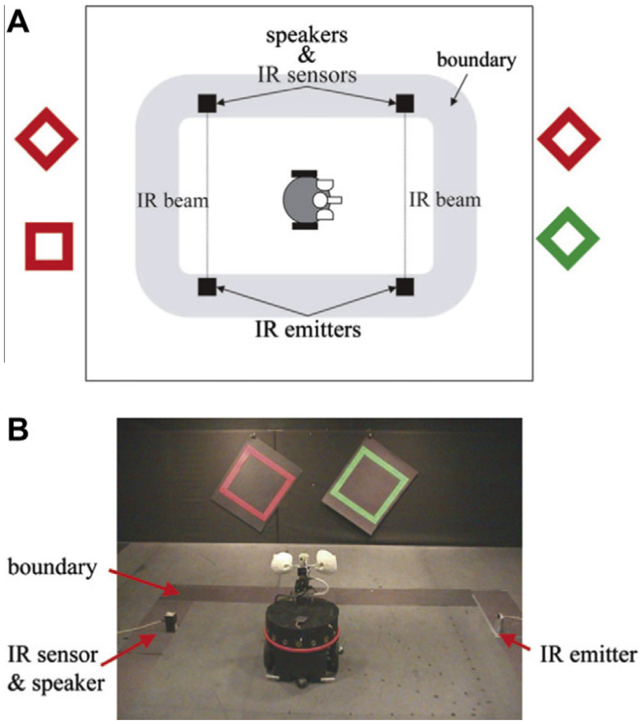

Fig. 1 Visual binding and categorization in a Brain-Based Device (BBD). (A) Experimental setup. Initially, an auditory cue signaled the brain-based device to orient towards a specific object category. After learning, the BBD oriented towards the object through vision alone. (B) Physical instantiation of the visual binding experiment. Adapted from Seth et al. (2004).

processing in the brain, then the robot might have whiskers. In a visual object recognition task (Fig. 1), we constructed a brain-based device that learned to attend to an object class in which feature binding was necessary for successful performance (Seth, McKinstry, Edelman, & Krichmar, 2004). Therefore, the agent not only needed to have a vision system that responded to features in its environment, but it also needed to demonstrate object recognition through a behavioral report (i.e., orienting and approaching behavior). In a somatosensory object recognition experiment (Fig. 2), we designed a whiskered robot that demonstrated its ability to discriminate different textures by freezing then escaping when it recognized a texture that was previously associated with a noxious stimulus (McKinstry et al., 2008). These types of tasks allow the behaviors of the robot to be compared with that of an animal model. Unfortunately, this approach can lead to ''one-trick ponies'', that is, the robot and its nervous system are designed for one specific function and cannot generalize to complete behaviors.

Agent design principle 2: the complete agent

When designing agents, one must consider the complete agent behaving in the real world.

In one sense, by building a robot and placing it in a physical environment, even if highly constrained, the system follows the complete agent principle. Many unplanned consequences emerge through interaction with the environment. For example, the act of moving through the environment and observing the world as a continuous stream of information can lead to invariant object recognition without the need for complex transformations (Krichmar & Edelman, 2002; Seth et al., 2004). Introducing saliency in the environment can lead to attentional signaling. For example, the robot CARL was designed to extract value from objects in its environment (Cox & Krichmar, 2009). Two out of four objects were salient to CARL and CARL learned the appropriate action for each (see Fig. 3). Unexpectedly, a strong attentional bias toward salient objects emerged through its experience in the real world, which could be observed both in CARL's behavior and in CARL's simulated brain.

On the other hand, designing a robot to test one specific aspect of cognitive behavior violates the complete agent behavior. As mentioned earlier, the community needs to

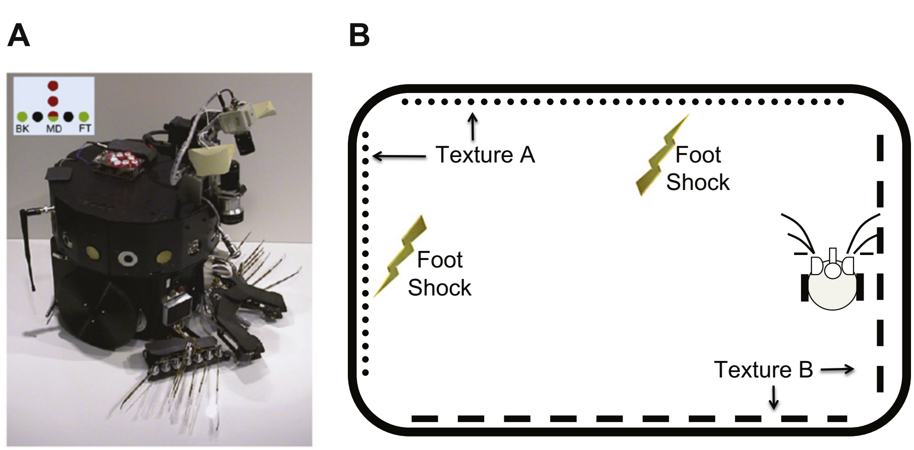

Fig. 2 (A) Darwin IX, a brain-based device (BBD) with whisker arrays. The arrangement of a whisker array is shown in the inset. Each array has 7 whiskers arranged in a row of 5 and a column of 3.The whiskers were made of two polyamide strips, placed back to back, that emitted a signal proportional to the bending of the strip. (B) Experimental setup for Darwin IX. The BBD explored a walled enclosure with textures Texture A and Texture B on the walls. Located on the floor adjacent to Texture A patterns were 'foot-shock' pads made of reflective construction paper. Darwin IX learned to either freeze or avoid the area where Texture A was experienced. Adapted from McKinstry et al. (2008).

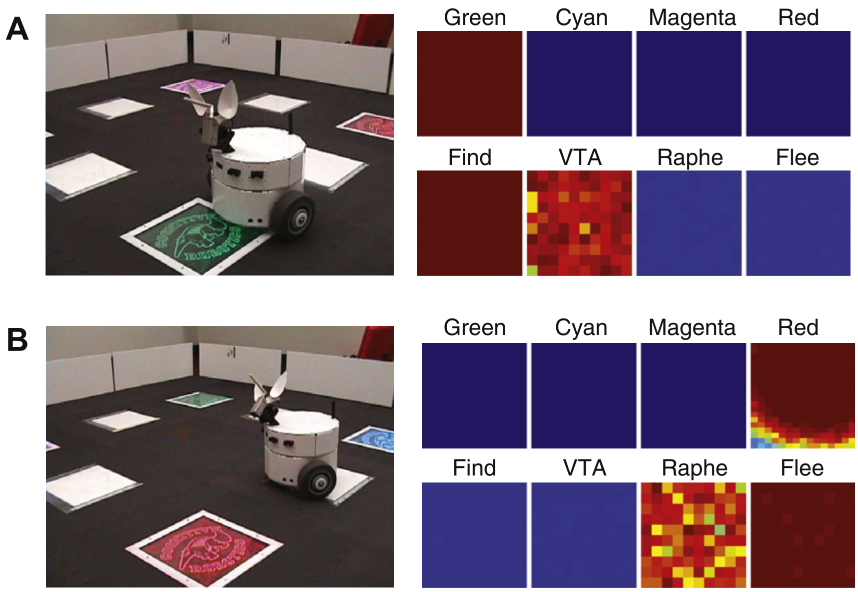

Fig. 3 CARL robot in colored panel task. The panels could flash any of 6 different colors. One color, green, signaled positive value. Another color, red, signaled negative value. The remaining colors were neutral. Signals were transmitted from the panel to a receiver on the bottom of CARL. (A) CARL during an approach or 'Find' response. The panels on the right show strong neuronal activity in its simulated green visual area, the dopaminergic system (VTA), and the Find motor neurons. (B) CARL during a withdrawal or 'Flee' response. The panels on the right show strong neuronal activity in its simulated red visual area, the serotonergic system (Raphe), and the Flee motor neurons. See Cox and Krichmar (2009) for more details.

show that their systems can have a suite of behaviors. Although cognitive scienctists and behavioral neuroscientists typically test their subjects in one task, it is understood that the subjects have a complete range of behaviors. In my own experience, there are two cases where our systems might have been considered complete agents. One was in RoboCup where we developed a robot to play soccer in the Segway Soccer exhibition league (Fig. 4). Although games such as these are limited by their rules, the robot did have to demonstrate a suite of behaviors (passing, shooting, navigating, positioning, strategic decisions, etc.) to be successful on the playing field (Fleischer et al., 2006). The other was a recently developed action selection system for robots based on principles of neuromodulation (Krichmar, 2012). Although simplistic in its design, this system can filter environmental events, and respond appropriately to the most urgent events. It is general purpose and can be tailored to the agent's niche and design (i.e., sensors, actuators, environment).

Agent design principle 3: cheap design

The construction and design of agents that are built to exploit properties of the ecological niche will be much easier or ''cheaper''.

The importance of this design principle can be observed when comparing the biped locomotion of passive dynamic walking robots to sophisticated humanoid robots, such as Honda's Asimo or Aldebaran's NAO. Passive dynamic walking robots exploit gravity, friction, and the forces generated by their swinging limbs (Collins, Ruina, Tedrake, & Wisse, 2005). As a result, they require very little energy or control

Fig. 4 (A) The Segway Soccer devices. On the left, a modified Segway HT scooter for the human player. On the right, a brain-based device (BBD) based on the Segway platform. (A) Active capture devices, (B) laser rangefinder, (C) pan-tilt unit and camera, (D) kicking assembly, (E) passive capture ring, (F) voice command module, and (G) crash bars. (B) The Segway Soccer BBD moves downfield through a cluttered playing field toward its opponent's goal. Adapted from Fleischer et al. (2006).

to move. These robots demonstrate what Pfeifer and Bongard call ''morphological computation'' in which processes are performed by the body and its exploitation of the environment, rather than by a central control system. In contrast, robots such as Asimo need complex control systems and long-lasting batteries to achieve the same result. I once visited Andy Ruina's Biorobotics and Locomotion laboratory at Cornell University and he shared with me that although he was not a biologist, he knew he had a good design when the mechanics looked natural and the energy expenditure was minimal. Exploiting the environment and energy minimization are hallmarks of biological systems.

Despite this basic principle in biological organisms, cognitive robots and their ilk rarely display cheap design. They typically require extensive computation, have very large power budgets, and tend not exploit properties of their niche. Some do exploit aspects of the environment; the invariant object recognition in an autonomously mobile robot described above is one specific example. But, in general, the complete cognitive agent is not designed with this in mind. An interesting exception is Barbara Webb's work with cricket phonotaxis (Webb & Scutt, 2000). In this system, a combination of a biologically plausible spiking neural network and the appropriate layout of auditory sensors led to a simple, yet elegant solution to phonotaxis.

From my own personal experience, our Segway Soccer team at The Neurosciences Institute solved a difficult sensorimotor problem with a very cheap design. On a fairly large playing field it was nearly impossible for our robot to catch a moving soccer ball given that it was large, cumbersome, and had a slow camera frame rate. Soccer balls would bounce off our robot before it had a chance to respond. Our team tested a variety of options to solve the problem of trapping a ball against the robot's body. Finally, we settled upon some cheap tubing that was fastened around the robot's body like a hula-hoop at just the right height (see Fig. 4). Any ball that was passed to the Segway robot was trapped by the tubing, giving the robot time to use its camera and proximity sensors to place the ball in its kicking apparatus. In a sense, this is what human players do when playing soccer. They use soft compliant materials angled appropriately to soften the impact of a ball coming toward them.

By putting more emphasis on designs that exploit the environment, we can offload some of the control from the cognitive robot's central nervous system onto the body itself. This should allow the robot to be more responsive to the environment and be more fluid its actions. In addition, it may free up the nervous system to put more emphasis on planning and prediction rather than reflexive movements.

Agent design principle 4: redundancy or degeneracy

Agents should be designed such that different subsystems function on the basis of different processes and there is overlap of functionality between subsystems. In this principle, Pfeifer and Bongard use the term ''redundancy'' to stress the importance of partially overlapping subsystems. I prefer the term ''degeneracy'' which is the ability of elements that are structurally different to perform the same function or yield the same output (Edelman & Gally, 2001). Degeneracy shows up throughout biology; from low-level processes such as the genetic code and protein folding to system-level processes such as behavioral repertoires or language.

A nice example of degeneracy at multiple levels is the Darwin X and XI brain-based devices, which were designed to demonstrate spatial and episodic memory (Fleischer & Krichmar, 2007; Fleischer et al., 2007; Krichmar et al., 2005). Darwin X solved a dry-variant of the Morris water maze and Darwin XI solved a place learning version of a standard plus maze (see Fig. 5). Both Darwins had an extensive model of the medial temporal lobe and its surrounding cortical regions. As these brain-based devices explored their environment, hippocampal place cells emerged.

Degeneracy occurred at the neuronal level

Because we were able to track every neuron in its simulated nervous system, we were able to trace the neuronal activity that led to hippocampal place activity. Although the CA1 place activity was similar on different trials when the brain-based device passed through the same location on the same heading, the neuronal activity leading to that neuron's place activity on a given trial differed dramatically.

Degeneracy occurred at the system level

Darwin XI received sensory input from its camera (vision), whiskers (somatosensory), compass (head direction), and laser range finder (depth/distance). Darwin XI's spatial memory was multimodal and degenerate. Even when one or more of its sensory modalities were lesioned, Darwin XI's behavior and place cell activity remained stable.

Degeneracy at the individual level

Nine different Darwin X subjects, which consisted of the same physical device but slightly different nervous systems due to variations in synaptic connection probabilities, solved the same spatial navigation task, but in unique ways. Some subjects bounced off the ''red'' wall to the hidden platform, some bounced off the ''blue'' wall, others went directly toward the platform location. The proficiency of each subject differed as well. Experience in the real world has a strong shaping effect on brain and behavior. For these reasons and similar to an animal experiment, we always run multiple subjects on a behavioral task. In every experiment I have worked on, no two brain-based devices or neurorobots have been alike, even when the nervous system and device were identical (Krichmar & Edelman, 2002).

Agent design principle 5: sensory-motor coordination

Embodied agents induce structured sensory stimulation through sensorimotor coordination by being situated in the environment and by manipulating the environment. For example, figure-ground separation can be achieved if a hand happens to push an object, resulting in the object moving with respect to its background (Fitzpatrick & Metta, 2003).

When our team at The Neurosciences Institute designed the visuomotor control for the Segway Soccer playing robot, we had to find a way to efficiently perform visual tracking and moving downfield with a heavy and somewhat sluggish two-wheeled Segway platform. Our approach was to develop a nimble pan and tilt system for its camera that quickly

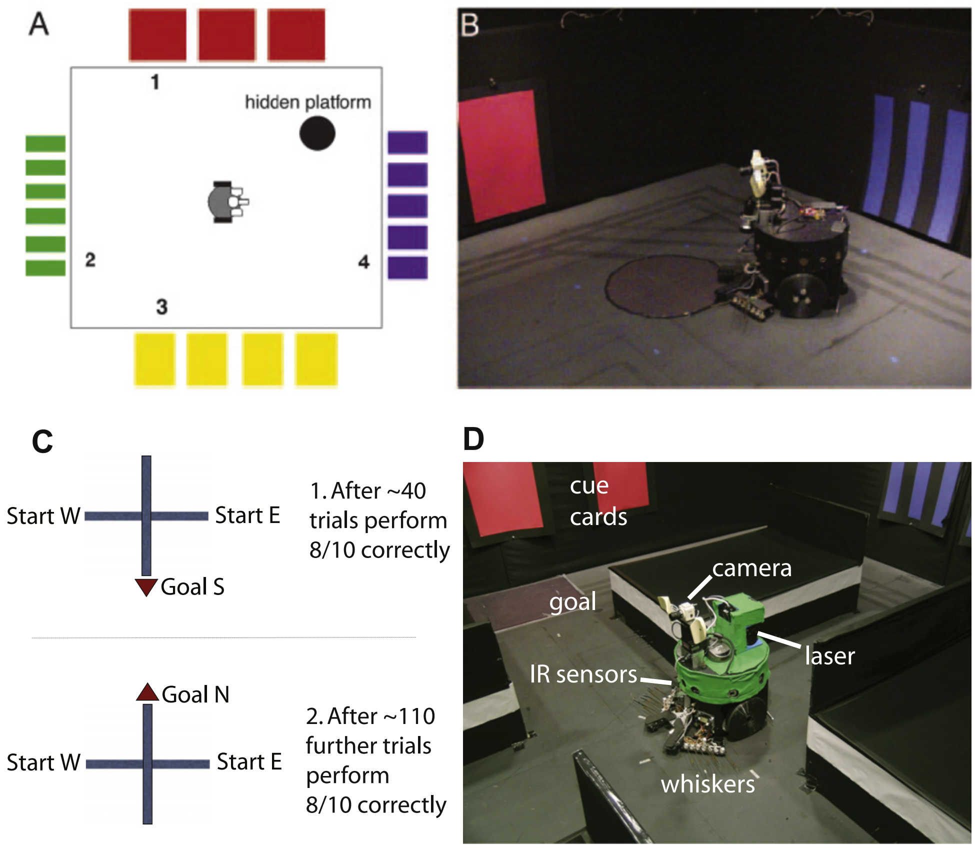

Fig. 5 (A) Experimental setup for Darwin X, a dry variant of the Morris water maze spatial memory task. (B) Physical instantiation of the spatial memory task. (C) Experimental setup for Darwin XI, a standard plus maze for place learning. (D) Physical instantiation of the plus maze. Adapted from Fleischer et al. (2007) and Krichmar et al. (2005).

saccaded to and foveated on salient objects on the playing field. The body followed where the camera pointed but at a slower rate. Specifically, if the camera was panned to the right, the Segway platform turned to the right in an attempt to center the camera. If the camera was tilted up because the object was further away, the Segway platform moved faster. As the Segway platform approached the object, its camera tilted downward causing the Segway platform to slow. The result was smooth, natural looking visual tracking with coordinated head and body movements (see Fig. 4 and http://www.vesicle.nsi.edu/nomad/segway/). This simple, yet elegant solution to visuomotor control was also implemented on our CARL robot for its operant conditioning tasks (Cox & Krichmar, 2009). The result is very natural looking and visitors to our lab tend to anthropomorphize CARL's behavior as it orients either toward a positive value object or away from a negative value object (see Fig. 3 and http://www.socsci.uci.edu/~jkrichma/CARL/robots.html).

Agent design principle 6: ecological balance

In a given task environment, there should be a balance between the complexity of the agent's sensory, motor, and neural systems. Moreover, there should be a balance between the morphology and its environment.

Of all of Pfeifer and Bongard's principles, designers of cognitive robots may violate this one the most. Most cognitive robots have very complex central processing or nervous systems, and a fairly simple robot. My work on the Darwin series of automata and the CARL robot tends to be guilty of violating this principle. Despite this imbalance, the embodiment has driven brain processing and led to results that would be difficult to imagine emerging in a computer simulation.

On the other hand, the behavior-based robotics community tends to be out of balance in the opposite direction. Many of the robot examples given by Pfeifer and Bongard have little or no neural processing. The field of evolutionary robotics evolves very simple neural controllers that bear little resemblance to real nervous systems (Floreano & Keller, 2010). The field of affective robotics is not necessarily interested in cognitive architectures or mechanisms, but instead are more interested people's responses to interactive robots with anthropomorphic designs and materials (Breazeal, 2004). Thus, these systems cannot demonstrate the executive control, planning, and learning typically associated with cognitive behavior.

I believe for cognitive robotics to move forward these disparate research communities need to come into balance. The materials, morphology, sensors, actuators, and the nervous system have to balanced and coordinated in their action.

There is another lingering issue, which has to do with ''How the body shapes the way we think''. Since the body is handling much of the computation in an intelligent agent, and cognitive processes are extremely slow to respond, so slow that it is a survival risk for the organism, then what is the brain good for? I believe the main functions of the central nervous system are predicting and planning for the future, and adaptation when the result does not meet expectations. The brain has numerous internal models that check the results of the agent's actions and make updates when there is a mismatch (Hickok, Houde, & Rong, 2011; Shadmehr & Krakauer, 2008). This is not a novel idea, but with respect to cognitive robotics and designing intelligent agents, it needs to be taken into consideration. The neural controller for the agent should be monitoring the body and peripheral nervous system of its body, and the agent's body itself needs to be capable of handling computation without direct neural control.

Agent design principle 7: parallel, loosely coupled processes

Intelligence is emergent from a large number of parallel processes, which are coordinated through embodied interaction with the environment.

The old way of thinking in cognitive science was ''sense, think, and act'', and this carried over to Artificial Intelligence robots. This way of thinking has changed in the computer science world in part due to the ubiquity of real-time and embedded systems. Most current computing devices, from phones to desktops, from onboard automotive computers to entertainment systems have parallel processes to handle asynchronous events. Designers of robotic systems typically use multitasking approaches to monitor the world concurrently. This allows them to interact with the environment asynchronously through multiple sensors and actuation systems. In this way, embedded systems are event-driven and multitasking. Similarly, biological organisms are event driven, multimodal, and have attention systems that are finely tuned to respond to novel information.

However, we tend to study cognitive science in a serialized fashion by focusing on one particular system at a time, be it a type of memory or a specific perceptual effect. The Cognitive Architecture community needs to move away from studying one system at a time, and rather study the complete agent, which responds to multiple, asynchronous events in a timely manner. One way to enforce parallel, loosely coupled processes might be through building physical systems having multimodal sensory systems, and embedding them in a dynamic and unforgiving environment (i.e., one that won't wait for the agent to think long and then act) that forces them to handle multiple processes concurrently.

Agent design principle 8: value

Intelligent agents are equipped with a value system that constitutes a basic assumption of what is good and bad for an agent.

Every brain-based device and neurorobot I have worked on had an innate value system that told the agent something was of intrinsic value and that triggered the appropriate reflexive behavior. The agent then learned which objects were predictive of value and tried to maximize the acquisition of good value and minimize the acquisition of bad value. Typically, these value-based robots employed models of the dopaminergic reward system to shape behavior.

Over the last few years we have expanded the simulation of value systems to include multiple neuromodulatory systems found in the vertebrate (Krichmar, 2008). Besides the dopaminergic reward system, there is the serotonergic system that is involved in harm aversion or the expected cost of an action (Cools, Roberts, & Robbins, 2008), there is the noradrenergic system that handles oddball or unexpected events (Yu & Dayan, 2005), and there is the cholinergic system that both increases and decreases the allocation of attentional resources (Baxter & Chiba, 1999). All of these systems are nuanced and they all interact with each other through direct and indirect pathways. All of these systems respond strongly to novelty, send broad signals to large areas of the cortex, and cause a change in network dynamics resulting in decisive action. Our first attempt at incorporating the interaction of multiple neuro modulatory systems into a cognitive robot controller has resulted in dynamic behavior that appears very much like a rodent exploring its environment (Krichmar, 2012).

One problem that remains unsolved in cognitive robotics is that these artificial value systems are dissociated from the agent's body. Real pain, hunger, thirst, and fatigue drive a true value system. Without this connection to bodily dependent drives, an artificial value system does not signal the immediacy of the agent's need and lacks to some degree the ability to shape the agent's behavior.

Conclusions

In the present paper, I have taken Pfeifer and Bongard's eight design principles thought to be necessary for an intelligent agent (Pfeifer & Bongard, 2007) and used them as a means to gauge the brain-based device, neurorobot, and cognitive robot approaches. Table 1 summarizes how well I believe the cognitive robotics community is following these principles and notes areas in which there is room for improvement. Although my views were based mostly on my own experiences, I feel this summary applies to the cognitive robotics community as a whole.

Note that up until this point I have spent very little time discussing the actual brain processing that leads to cognition. In no way am I arguing that this is unimportant. It is essential that an intelligent agent have some brain control. I would further argue that a necessary condition for studying cognition and building a cognitive artifact is that its artificial brain must resemble its real counterpart (Krichmar & Edelman, 2005; Krichmar & Wagatsuma, 2011). However, this alone is not sufficient; we need to bring the brain, body, and environment into balance in our designs.

Where are we failing?

In general, the community is failing to make systems that perform more than one function at a time. Most of these systems exist in sterile, highly controlled laboratory settings. Most of these systems rely too much on neural control driving the body and behavior instead of the other way around. I would also add from a neuroscience standpoint,

Table 1 How well cognitive robots follow Pfeifer and Bongard's agent design principles.

Design principle	Adherence to principle	Room for improvement
Three constituents principle	Yes, in laboratory settings	Need to demonstrate cognitive capabilities outside the lab
Complete agent	Yes, when embodied No, because too specialized	Need more general-purpose systems
Cheap design	Very rare	Cognitive robots depend too much on top-down brain processing. They need to be designed to exploit the environment and listen to their body
Degeneracy	Yes	Cognitive robots should be designed with degeneracy in mind. Observers should take note of how their system displays degeneracy
Sensory-motor coordination	Yes	Cognitive robots usually follow this principle, but it could be exploited by designing more interesting behaviors
Ecological balance	No	Need to put more emphasis on the body handling processing (i.e. morphological computation) to offload top-down brain control
Parallel, loosely coupled systems	Yes, in practice No, in the analysis and study	Most cognitive robotic systems follow this principle in their design. But, the behaviors they exhibit do not exploit parallelism and concurrency
Value	Yes, in artificial value systems	For value to have true meaning, it needs to directly affect the agent's body

most of these systems are focused on cortical processing, and tend to ignore critical subcortical processes. However, this is a reflection of neuroscience community in general, and not just the modelers.

We need to put more emphasis on the morphology of our cognitive robotic systems. Brains do not work in isolation; they are closely coupled with the body acting in its environment. *The brain is embodied and the body is embedded in the environment*. However, it goes beyond that. Biological organisms perform *morphological computation*, that is, certain processes are performed by the body that would otherwise be performed by the brain (Pfeifer & Bongard, 2007). This would allow the central nervous system, which is slower and requires more processing, to predict, plan, and adapt by comparing its internal models with current information from the body (Hickok et al., 2011; Shadmehr & Krakauer, 2008).

Why is this important?

Putting the brain, body, and environment into balance is critical for several reasons. First, having this balance would allow the brain processing of the agent to be more closely coupled with the body of the agent, and in turn its interaction with the environment. In biological organisms, the nervous system, which includes the spinal cord, subcortical regions, and the cortex is in close interaction with not only a wide array of external sensors, but also internal sensors that are monitoring all aspects of the body (Damasio, 1994). In a sense it is very difficult to separate the brain from the body. In fact, it is nonsensical since the brain is technically part of the body! Second, having this balance will allow the body to handle many of the reflexive, rapid movements and responses to environmental events, and allow the central nervous system to handle what has been traditionally been called cognitive. That is, planning, executive control, predicting future outcomes, decision-making, and adapting to improve future actions. Third, cognition requires action and feedback from its body. The action may be outward, or it may be internalized. However, the end result is a future action. Planning for a future action requires some sense of the effect the action will have on the body. If it is a forward model, there is a prediction of what the agent should sense after an action. If it is an inverse model, there is a prediction of the body state after an action. Even high-level planning over a long time period needs to be grounded at some level to physical action. Fourth, having this balance should allow for the construction of more flexible and general-purpose agents. These complete agents could then demonstrate a wide range of capabilities over a broader set of environments. Such an agent might be more comparable with a biological organism, might be considered truly cognitive, and may have some societal benefits as well.

Following up on that final point, the BICA community needs to demonstrate that cognitive robots are practical. Although I have personally gained a better understanding of how the brain works by building these artifacts, it may not be enough. It is not just that the BICA community has to show that these systems are practical to gain legitimacy; they also need to show that something is to be gained by taking a cognitive approach. Recent technological advances in affordable computation, sophisticated sensors, and rapid prototyping give us an immediate window of opportunity. Building artifacts that demonstrate a wide range of cognitive behaviors in real, physical environments would be a step in that direction. To do so will require having the brain and body in register.

Acknowledgments

The author would like to thank the Cognitive Anteater Robotics Laboratory (CARL) at the University of California, Irvine, his Cognitive Robotics students, Hiroaki Wagatsuma of the Kyushu Institute of Technology, and the ''Build a Brain Team'' at The Neurosciences Institute for many inspiring discussions. Supported by the National Science Foundation (Award Nos.: EMT/BSSE-0829752 & IIS/RI-0910710) and the Office of Naval Research (Award No.: N000140910036).

References

Arkin, R. C. (1998). *Behavior-based robotics (intelligent robotics and autonomous agents)*. Cambridge: The MIT Press.

Baxter, M. G., & Chiba, A. A. (1999). Cognitive functions of the basal forebrain. *Current Opinion in Neurobiology, 9*(2), 178–183.

Bongard, J., Zykov, V., & Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science, 314*(5802), 1118–1121.

Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. Cambridge: MIT Press.

Breazeal, C. L. (2004). *Designing sociable robots*. Cambridge, MA: The MIT Press, A Bradford Book.

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

Clark, A., & Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior, 7*(1), 5–16.

Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science, 307*(5712), 1082–1085.

Cools, R., Roberts, A. C., & Robbins, T. W. (2008). Serotoninergic regulation of emotional and behavioural control processes. *Trends in Cognitive Sciences, 12*(1), 31–40.

Cordeschi, R. (2002). *The discovery of the artificial: Behavior, mind, and machines before and beyond cybernetics*. Kluwer Academic Publishers.

Cox, B. R., & Krichmar, J. L. (2009). Neuromodulation as a robot controller: A brain inspired design strategy for controlling autonomous robots. *IEEE Robotics & Automation Magazine, 16*(3), 72–80.

Craik, K. J. W. (1967). *The nature of explanation*. Cambridge University Press.

Damasio, A. (1994). *Descartes' error Emotion reason and the human brain*. New York, NY: Penguin Books.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America, 98*(24), 13763–13768.

Edelman, G. M., Reeke, G. N., Jr., Gall, W. E., Tononi, G., Williams, D., & Sporns, O. (1992). Synthetic neural modeling applied to a real-world artifact. *Proceedings of the National Academy of Sciences of the United States of America, 89*(15), 7267–7271.

Fitzpatrick, P., & Metta, G. (2003). Grounding vision through experimental manipulation. *Philosophical Transactions A, Mathematical, Physical, and Engineering Sciences, 361*(1811), 2165–2185. http://dx.doi.org/10.1098/rsta.2003.1251 [Research Support, Non-U.S. Gov't].

Fleischer, J. G., Gally, J. A., Edelman, G. M., & Krichmar, J. L. (2007). Retrospective and prospective responses arising in a modeled hippocampus during maze navigation by a brain-based device. *Proceedings of the National Academy of Sciences of the United States of America, 104*(9), 3556–3561.

Fleischer, J. G., & Krichmar, J. L. (2007). Sensory integration and remapping in a model of the medial temporal lobe during maze navigation by a brain-based device. *Journal of Integrative Neuroscience, 6*(3), 403–431.

Fleischer, J. G., Szatmary, B., Hutson, D., Moore, D. A., Snook, J. A., Edelman, G. M., et al. (2006). A neurally controlled robot competes and cooperates with humans in Segway soccer. *Paper presented at the IEEE international conference on robotics and automation*, Orlando, FL.

Floreano, D., & Keller, L. (2010). Evolution of adaptive behaviour in robots by means of Darwinian selection. *PLoS Biology, 8*(1), e1000292.

Grey Walter, W. (1953). *The living brain* (2nd ed.). London: Penguin.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron, 69*(3), 407–422.

Iida, F., & Pfeifer, R. (2004). Self-stabilization and behavioral diversity of embodied adaptive locomotion. In F. Iida, R. Pfeifer, L. Steels, & Y. Kuniyoshi (Eds.), *Embodied artificial intelligence* (pp. 119–129). Berlin: Springer.

Krichmar, J. L. (2008). The neuromodulatory system – A framework for survival and adaptive behavior in a challenging world. *Adaptive Behavior, 16*, 385–399.

Krichmar, J. L. (2012). A biologically inspired action selection algorithm based on principles of neuromodulation. Paper presented at the international joint conference on neural networks, Brisbane, Australia.

Krichmar, J. L., & Wagatsuma, H. (2011). Neuromorphic and brain-based robots. In A. V. Samsonovich & K. R. Jóhannsdóttir (Eds.), *Biologically inspired cognitive architectures* (Vol. 233, pp. 209–214). IOS Press.

Krichmar, J. L., & Edelman, G. M. (2002). Machine psychology: autonomous behavior, perceptual categorization, and conditioning in a brain-based device. *Cerebral Cortex, 12*, 818–830.

Krichmar, J. L., & Edelman, G. M. (2005). Brain-based devices for the study of nervous systems and the development of intelligent machines. *Artificial Life, 11*(1–2), 63–78.

Krichmar, J. L., Nitz, D. A., Gally, J. A., & Edelman, G. M. (2005). Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task. *Proceedings of the National Academy of Sciences of the United States of America, 102*(6), 2111–2116.

McKinstry, J. L., Seth, A. K., Edelman, G. M., & Krichmar, J. L. (2008). Embodied models of delayed neural responses: spatiotemporal categorization and predictive motor control in brain based devices. *Neural Networks, 21*, 553–561.

Pfeifer, R., & Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge: MIT Press.

Seth, A. K., McKinstry, J. L., Edelman, G. M., & Krichmar, J. L. (2004). Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device. *Cerebral Cortex, 14*(11), 1185–1199.

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research, 185*(3), 359–381.

Webb, B., & Scutt, T. (2000). A simple latency-dependent spiking-neuron model of cricket phonotaxis. *Biological Cybernetics, 82*(3), 247–269.

Wyeth, G., & Milford, M. (2009). Spatial cognition for robots robot navigation from biological inspiration. *IEEE Robotics & Automation Magazine, 16*(3), 24–32.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron, 46*(4), 681–692.