Autonomous Robots with Both Body and Behavior Self-Knowledge

B Brent Gordon
NASA Goddard Space Flight Center
Information Science and Technology Research Group
Greenbelt, MD 20771 USA
bgordon@backserv.gsfc.nasa.gov

Abstract — At PerMIS'06 Gage suggested that it should be standard practice for autonomous robots to log all "sensory inputs, internal states, and behaviors," but all the benefits he derived from doing so depended on humans interpreting those data. In this paper we ask what value the robot might find if those data were available to it in real-time? This question leads us first to a careful review of the human somatosensory system, and then to an architecture of multiple paired forward and inverse motor control subsystem and sensorimotor context specific models in the cerebellum, as plausible biological inspirations for robotic internal self-models. Based on these, the paper describes a *robotic proprioception modeler* (rpm) whose most striking potential application is to allow an autonomous behavior-based robot to monitor for and detect anomalous plant or control behavior. With the addition of a little cognitive ability, the rpm should be able to assess and resolve anomalies as well.

Keywords: *robot, autonomous robots, internal models, self-knowledge, behavior-based robotics, biologically inspired robotics, sensorimotor control.*

I. Introduction

The proximate cause for this paper's existence is D. Gage's presentation at last year's (2006) PerMIS [23], in which he advocated that it should be a standard design feature for an autonomous robot to "log 'all' its sensory inputs, internal states, and behaviors." His argument was that doing so would "facilitate continuous system improvement and greatly simplify the evaluation process," as well as simplify debugging during development, and thus the up-front costs would be outweighed by the overall benefits. That left me wondering from an AI perspective, what an autonomous robot could do with all those data, i.e., what sort of benefit and value it might be able to gain.

In order to get some insight into the problem, I decided I would look into the biological analogy. Therefore I asked:

1) What do these data Gage suggests collecting correspond to in humans?
2) How do humans use and take advantage of them?
3) Could an autonomous robot use the data Gage suggests collecting to its advantage in analogous ways, and what would those analogies look like?

The one thing about question 1 that is not so straightforward is that it actually has two (or more) parts, one having to do with sensing of the world, of the body and of where the two come together, and the other part having to do with collecting information about and modeling the motor control system. While question 2 is the key, the fact that, as is often the case for medical matter, much of the evidence comes from the absence of a sense or model creating a difficulty, and this is harder to justify as an analogy. As for question 3, and biological analogies in general, one must be careful to find the biological principles involved, and a good, meaningful robotic analogy, as opposed to attempting to copy a biological structure, unless that was the intention in the first place.

An interesting, and slightly unexpected, connection with this year's Workshop Theme, *the interplay between autonomy and intelligence,* our investigation turned up was that adding enough of the right self-knowledge about its physical and motor control system to an autonomous (behavior-based) robot that didn't have explicit models of those things could improve the robustness of its autonomy, but only if there was also enough intelligence to use that knowledge effectively. From an evolutionary perspective, one might easily imagine the cycle of more robust autonomy, more self-knowledge going into a more complex modeling and management process, more intelligence required to use it more effectively, as a spiral.

The next section describes the human somatosensory system and a certain model of the sensorimotor control system, while the section after that runs quickly through a catalog of how those models are put to good use. Then we turn to robots as section IV sets up the construction of the robot proprioception modeler (rpm) and the next section after that describes how autonomous robots might put it to good use, especially for anomaly detection and, if there is some cognitive capability hanging around, resolution. Having finally discovered what this paper is about, we can tell which papers are related, and then draw some conclusions.

II. Human sensing and sensorimotor models

The first thing we do is review the full range of human external, by which we mean environmental, and internal, or self-, sensing, breaking it down along functional lines. Our classification is based on the sense's primary application. Then we briefly look at an architecture containing multiple pairs of predictive and generative models for components of the motor system. This architecture plays the role of a complex, context-sensitive model of the entire motor system. Taken together those sensory data plus the data of the architecture seem to be a good analogy to the data Gage suggests collecting. We conclude the section by indicating some ways that humans use models based on these data, for an answer question 2.

A. The nervous system

First we need a few words about the nervous system, which plays a crucial role in all this. Recall that the *central nervous system* consists of the brain and the spinal cord; all the rest belongs to the *peripheral nervous system.* Nerves in the peripheral nervous system that take signals toward the central nervous system are referred to as *afferent;* these are the sensory ones with which we will be mostly concerned. The thing to be mindful of is that their far ends take on numerous forms as specialized *receptors,* which may perform a sensory function themselves or may act as interpreters between a sensory organ and the nervous system—we'll see examples of both in a moment. Nerves that carry signals out from the central nervous system, the activation or motor nerves, may also be referred to as *efferent.* Efferent peripheral nerves are classified as *autonomic* or *somatic,* where the former are those that activate or inhibit internal organs and the latter are those that activate the skeleto-muscular system. All afferent nerves are considered somatic, including those sensing things in or about internal organs or coming from the sensory organs on the head.

B. Four external senses

In the class of *external* senses we include vision, hearing, smell, and taste, since the primary function of each of these is to discern some property or something about the nature of the world around us. Only vision is a special case, since perhaps as much as 40% of its capacity is devoted to proprioceptive purposes [8].

C. The vestibular and haptic senses

The *vestibular system* is a sensitive and elaborate inertial sensor system. Structurally it consists of three mutually perpendicular *semi-circular canals* and two linear components called *otoliths,* all of which are lined with hair-like receptors with a variety of sensitivities, including, for example, slow acceleration, faster acceleration, and even jerk, the derivative of acceleration [8]. Under the assumption of a constant inertial field, for example, if you are earthbound, it would be a legitimate design decision to treat the vestibular sense as if it were an internal sense that provides feedback about body—well, head—motion and orientation.

Although the *sense of touch* is one of the five senses we learn about as children, it really breaks down into four distinct basic kinds of touch: cold, heat, pressure, and pain. But the situation is even more complicated than that, because the many, many different receptors distributed unevenly over the body respond to different levels and combinations of these.

D. The kinesthetic sense

For the purposes of this paper, a person's *kinesthetic sense* is the sense that provides the person with knowledge of where the parts of that person's own body are and how they are moving. The corresponding *kinesthetic receptors* specifically consist of *joint receptors* embedded in the joints, the *Golgi tendon organs* embedded in the tendons, *neuromuscular spindles* embedded in muscle fibers, and *cutaneous receptors* embedded in sub-dermal cutaneous tissue; where the latter three are sensitive stretch [8]. What we are calling the kinesthetic sense here is often referred to as "proprioception," but the latter term is ambiguously used for both proprioceptive sensing and proprioceptive awareness [24], and apparently is not or is assumed to include the vestibular receptors depending on which side of the Atlantic you speak English [20]. So with all that baggage and ambiguity, we will just use the term "kinesthetic sense."

Now, to emphasize how the kinesthetic sense is like most of the other senses, except for its receptors being internal and distributed, consider an analogy with the eye. It has a very large number of retinal sensors and a very large number of nerve fibers carrying signals to the brain, similarly as the kinesthetic sense, except that its sensors are all located spatially close together. Another point of contrast is that the eyes normally impinge on attentive consciousness although we can choose to turn them off, whereas the kinesthetic sense normally stays in the background and comes to the foreground only when called. Think of your right foot right now.

Remark: This result from one recent study concerning kinesthetic sensing seems interesting. "Neuroimaging studies (PET, fMRI, TMS) have clearly demonstrated that somatic perception of limb movements engages the human motor cortex . . . with neither overt limb movements nor intention of movements" [38]. One plausible speculation from the paper is that the motor cortex is a part of the network in the brain that is responsible for updating its internal models; *cf.* [3]

E. Non-kinesthetic internal sensing

While for finding biological lessons that might be transferable to robotics it seems reasonable to separate the internal kinesthetic sensing from the rest of the internal sensing, as a non-expert in anatomy or neurology I don't know a good technical name for what is left over. Very approximately, it would be the part of the afferent, somatic, nervous system that corresponds to the (efferent) autonomic nervous system. In any case, without attempting to list all the components, since we don't need that level of detail, what we want *non-kinesthetic internal sensing* to refer to is all the receptors throughout the body that sense all the states and conditions pertaining to the internal organs, glands, etc., for example hormone levels, and how full the bladder is, as well as all the states and conditions that have to do with homeostasis, for example, glucose, salt, potassium levels, plus anything else that should be included.

Table I summarizes the terminology we have introduced in this section.

TABLE I
SOMATIC SENSORY (AFFERENT) NERVOUS SYSTEM

External senses	Vision Hearing Smell Taste
Vestibular sensors	3 semicircular canals 2 otoliths
Haptic (touch) basic senses	Cold Heat Pressure Pain
Kinesthetic receptors	Joint receptors Golgi tendon organs Neuromuscular spindles Cutaneous receptors
Non-kinesthetic internal (states and conditions of)	Internal organs Cardio-vascular system Endocrine system Gastro-intestinal system And the rest

F. An architecture of multiple forward-inverse model pairs

The question was whether or how humans model their motor control subsystem, in the context of the overall sensorimotor system, and what [11] [47] [46] [29] [27] [30] proposed, and evidence supports, is the idea that the cerebellum contains a very large collection of coupled inverse pairs of models, i.e., each pair consists of a forward, or predictive, model and its corresponding inverse, or generative, or controller, model. Moreover, this collection forms an architecture, in the sense that there is some mechanism for properly combining model pairs to produce a desired behavior, for example.

To be more specific, and perhaps clearer, the idea is that one model pair models one specific motor subsystem in one specific sensorimotor context, for example, the arm bent at some particular angle while that hand holds approximately some particular weight [11]. How wide a context might be, or how narrow it must be, is difficult if not impossible to tell from the experiments that have been done so far, so take this with a grain of salt. And although we can imagine that there might be infinitely many contexts, in practice a new motor subsystem-context model pair is not learned and added to the collection until it is actually needed ("lazy learning").

Now, suppose we have a behavior to be executed, and some number of model pairs are already in the architecture. Then the basic idea is that after each time step the predictions made by each forward model can be compared with the sensory feedback, and the level of involvement of that model pair adjusted accordingly. In particular, the inverse model is incorporated at the appropriate level in generating the next control command. If the existing model pairs do not account for everything in the behavior to be executed, then there is something new to learn. Of course the actual process must be somewhat more complicated than that, but this captures the concept sufficiently for now.

Remark: At the time of writing I have not yet worked out the relationship, if any, between this model and the cerebellar models of Albus [1] and Marr [34]. So that is left as an exercise for the reader.

III. HOW HUMANS USE THESE SELF-MODELS

Now we consider question 2. The data available to a human consist of the somatosensory data described in II-B, II-C, II-D, II-E, plus the multiple model pairs, as data themselves plus whatever data they produce, described in II-F. With all these, what can a person do? Throughout the discussion we will implicitly assume that these data influence the human, or that the human brain applies them, through some sort(s) of model(s). It is possible that understanding more about the mechanisms could provide useful inspiration for robotic implementation, but that will have to wait for another time. For now those details need not concern us. The following discussion has some overlap with the discussion of internal models in [26].

A. Improving sensorimotor control

The first kind of application has to do with improving sensorimotor control. Actually, as humans are designed, it is possible, but only with enormous effort, to get by with visual (and vestibular) proprioception alone, that is, without kinesthetic feedback due to disease [20]. Even when kinesthetic sensing is artificially inhibited (by vibration) in one limb, otherwise healthy subjects found it more difficult to adapt to tasks mediated by unusual forces or that required bimanual coordination [39] [31] [42] [45] [10] [9] [32] [33] [7]. This is consistent with the cerebellum's known role in mediating fine motor control [18] [40]. We won't go into detail here now, but predictive models are useful in compensating for high latency in sensory feedback, such as is the case with kinesthetic sensing [36] [45]. And we refer to [47] [46] for discussions of the advantages of paired forward-inverse models over either one alone.

B. Better sensorimotor simulations

The level of detail is there to run much rich and sophisticated, accurate, and lengthy in the sense of projecting further into the future, sensorimotor simulations. Humans use this ability at many time scales and levels of detail, from running through ones mind the possible next moves, to estimating if one will fit through that opening over there, to trying to improve one's golf game by visualizing good technique.

C. Creating behavioral habits

While a person is learning a new motor skill such as walking or riding a bicycle, at first it tends to demand all of the person's conscious attention, but eventually the learning consolidates, a habit is formed, and, depending on the activity and conditions, it may seem to require very little or no conscious attention. Presumably the necessary model pairs have been created and added to the architecture in the cerebellum, and something else that represents the necessary sequencing and hierarchical structuring of model pairs into a behavior like pedaling a bicycle, which involves numerous motor subsystems and sensorimotor contexts. Our point here, though, is that wherever a behavior is learned and consolidated into a habit, it frees resources for new learning [9].

D. Anomaly monitoring and detection

Without a doubt the most important application for such a complete and detailed model of the human physical and sensorimotor control system is anomaly detection. As a first layer of monitoring, the haptic, kinetic, and internal non-kinesthetic sensory systems, by virtue of extending so thoroughly throughout the body, are pretty good at detecting and providing a first warning of most problems. But at a next level, the model expects vision and vestibular, haptic, and kinesthetic sensing all to correlate [8] [42]. And it is something at a higher level still, probably involving the architecture of multiple model pairs but also probably much more, when neurological malfunctions, say, manifest themselves through anomalous sensorimotor control or behavior.

E. Metacognition

Humans are capable of *metacognition,* that is, thinking about their thoughts and modeling their models, *cf.* [37]. And while it would be too much to consider this in any great generality, one particular formulation of metacognition, namely the "metacognitive loop" of [4] [5], will be useful to us. Briefly and simply, the key points of the metacognitive loop are that it should "*notice* when something is amiss, *assess* the anomaly, and *guide* a solution into place," where an anomaly is defined as "a deviation from expected values or outcomes" [4]. And we have just seen in the previous paragraph (III-D) that the availability of complete somatosensory data enables anomaly detection. As for assessment, the very detailed level of simulation that are possible don't guarantee success, but provides strong support to whatever cognitive system is overseeing the assessment. And as for guiding a solution into place, that would harken back to improving control (III-A) or creating a new habit (III-C). Indeed, we suspect that when the metacognitive loop is only "meta" with respect to the sensorimotor control system, most of its operation takes place at a sub-attentive if not subconscious or unconscious level. And since the sensorimotor control system, at least the aspects of it we are discussing, we don't think of as conscious, maybe the metacognitive loop would better be thought of as a manager or meta-manager at this level?

IV. Robotic sensing and self-modeling

Now we return to the original notion of collecting and logging "all sensory inputs, internal states, and behaviors" [23] produced by an autonomous robot, but now with some feeling for what we might want those data to support in the way of models and processes. Before we can get to that, however, we have to choose an autonomous robot architecture to work with, make sure our system is outfitted with all the sensing the problem calls for, and set up a process to set up all the necessary models.

A. Autonomous behavior-based robot

1) Robot: Just to be sure we are all thinking about the same thing, for the purposes of discussion for the rest of the paper a *robot* refers to a constructed, embodied, and situated system, where a system is *embodied and situated* if it physically exists in some actual environment somewhere, and it is able to both sense that environment as well as initiate interaction with it [41] [48] [2] [19].

2) Behavior-based architecture: In the spirit of scientific inquiry, implicitly in our investigation, we would like to understand as well as possible how much and what kind of difference it is making if we do this or that with the data Gage suggests collecting, or with the internal models they implicitly determine, or require. What I'm trying to get at here is comparisons would be made easier if we started with a robot architecture that incorporated no internal models; and the most natural choice of architecture given that request is the *behavior-based* architecture [6] [35]. Following [6], the basic principles of behavior-based architecture are:

- "emphasis on a tight coupling between sensing and action;"
- "decomposition into contextually meaningful units (behaviors or situation-action pairs);"
- "aversion to the use of representational [symbolic] knowledge."

The original and most basic example of a behavior-based architecture is the "subsumption" architecture introduced by Brooks [16]. It avoids representational knowledge and provides a tight coupling between sensing and action by virtue of being a motor control system, without any adaptivity built into it; behavioral decomposition is a non-trivial design feature. Most behavior-based systems have some learning or adaptivity built into them, or some other feature that disqualifies them from being instantiations of the subsumption architecture, see [6] for many examples.

Continuing in the spirit of scientific inquiry, we will assume that our test robot starts out with some "reasonable and normal" level of self-sensing, for example, we would expect all wheels to have encoders and all arm joints to have torque or force and joint angle or distance sensors, as appropriate.

3) Autonomy and intelligence: Finally, the last major assumption we are making, just because it is the case we are interested in—and it will turn out to be quite interesting, too—is that the robotic system we work with is "autonomous." We really don't have a good definition for autonomy, except that for the purposes of this paper we can distinguish it from intelligence by saying that *autonomy* has more to do with behavior and acting independently while *intelligence* has more to do with an ability to exploit information or knowledge advantageously. Beyond that, we will let your intuition about what they mean suffice, except to remind you that robots come at all different levels of autonomy, or at least the word is used with many different meanings.

B. Additional internal sensing

Even if the original question might have been about collecting all of a robot's sensory, state and behavior data, since looking at the human somatosensory system and how humans it, an unstated but implied requirement that the level of internal sensing be increased from "reasonable and normal" to "thorough" has become apparent. In other words, it may be that the first thing that has to be done is to add more internal sensors to the subject robot. What kind? Most likely a large number that could be considered analogous to non-kinesthetic internal sensors, since in all likelihood whatever external sensors are present suffice, since we are assuming that the system is autonomous, and as described above, a reasonable and normal level of sensors typically covers most if not all degrees of freedom. For example, the additional sensors might include such things as power levels, temperatures, resistance to movement, lubrication levels where relevant, resistance in electrical circuitry, wherever anything can be measured or wherever anything could change or go wrong.

In addition to the extra sensing, there is a software and firmware issue as well. In reality, one doesn't want to *add* the kind of somatosensory system with sensory, state, and behavior data collection and modeling we are looking at here to an already complete system, it should be integrated in by design from the start. However, for a thought-experiment, it is conceptually cleaner to start with the autonomous system and then add the extra sensing, and now change the necessary software and/or firmware to report all the control system elements' states and behaviors, as well.

C. The robot proprioception modeler

At this point we need to start setting up the modeling system that we will use, for which we have chosen the name **r**obot **p**roprioception **m**odeler, or **rpm**. The idea is that the **rpm** will receive, or be granted access to, the data coming in from all the external and internal sensors as well as that software or firmware that is reporting on the states and behaviors of the control system's element, all in real-time, even as all the same data is being logged. In the meanwhile, there are two models we have yet to deal with creating, but once that is done, it will also be up to the **rpm** to maintain those models as well.

One of these models should be the analog on the robotic side of the architecture of multiple pairs of forward and inverse models on the human side, but we are not entirely sure what exactly will play this role. My suspicion is that the way to think about the way to make an analogy is to think of the behavior-based control system as if it were some analogy of the brain, issuing commands to move in various ways for reasons that are inaccessible to the **rpm**. Nonetheless, the **rpm** can track the commands, their contexts, the responses, etc. The other model should be some sort of complex-systems multi-layer, multi-resolution, multi-perspective model of the robotic system as a physical entity in a physical world. How fine a granularity can be achieved depends on the distribution and resolution of the internal sensors, noting that perhaps more can be inferred than can be directly sensed.

For both these models, maintaining them once they exist in some reasonable form is likely to be much easier than creating either in the first place. The most critical principle here is to make sure that the models are ultimately assembled accurately from (accurate) sensor and control system data. Within the bounds this creates, my inclination would be to look for what arrangement of telling the robot its specifications and having it revise and correct those to what it senses, for example, versus having it learn for itself from scratch, would work most effectively, *cf.* [22] [12] [15].

V. Anomaly detection and management

While we don't want to omit a few other worthwhile benefits that an autonomous behavior-based robot can get come from these data and the internal models that go with them, by far the most significant benefit is the potential for anomaly monitoring and detection and the consequent potential for intelligent robustification of the system, in the sense of making it much more perturbation tolerant. This does not come with the models for free, however; the models supply a suitable tool, but the intelligence to use it must still be supplied, somehow. Nonetheless, we get a very sharp image of the interplay between autonomy and intelligence, as well as a clear example of the robotic mind-body problem.

A. Better sensorimotor control

While there are some who argue that biological motor controllers need forward models in the loop because of the long latency of the sensory feedback, even robotic systems suffer some latency, and the problem may become more acute before it gets better, *if* the need to interpret raw sensory data as perception grows, as it is likely to. With more thorough kinesthetic sensing, and, let's suppose, context-sensitive linked pairs of forward and inverse models for each motor system, the constraints are sufficient that the control problem is easier, not harder, and still the process may allow for finer control (subject to hardware limits, eventually).

The whole-body model might simplify the coordination of movements whose coordination was not originally built into the behavior-based architecture.

B. Better sensorimotor simulations

With the more extensive and comprehensive models that come from more systematic and complete sensing, better simulations are possible. The availability of sufficient computing power to take advantage of all the additional detail is always a question, but an alternative is to be selective about what parts or aspects of the model to use, too. The point here is that more information creates options.

C. Non-kinesthetic factors

Since a lot of non-kinesthetic sensing is intended to be part of these models from the start, it might be easier or more natural to create connections between behavior and other kinds of factors. For example, power-aware movement, or temperature-aware might be useful in some circumstances.

D. Anomaly detection

Autonomous behavior-based robots detecting faults or anomalous behavior in themselves is a much more significant application of the **rpm** than the other applications mentioned so far. Although I have not, at the time of this writing, looked into this thoroughly, self-detection of anomalous behavior in autonomous robots is not a topic that seems to have received a great deal of specific attention, though there has been some, and of course general systems principles apply. In any case, similarly as for the human case, the first layer of monitoring corresponds to the expanded internal sensing system that was installed as part of this process. Then at a next level, the multi-layer body model could reasonably be expected to include any number of correlations between and among kinesthetic and non-kinesthetic receptor inputs, meaning, in particular, that there is an idea of what a normal pattern is, and thus what an abnormal pattern is. I am inclined to think that the problem becomes much more complex when the robotic analog of the architecture of multiple pairs of linked forward and inverse models is included in the picture, but for the reasons mentioned in (V-A), namely that the number of constraints grows as well, it may not. In any case, it should be possible at least under some circumstances to detect anomalies at the control, as compared to the plant, level.

E. Anomaly handling requires cognition

Once the **rpm** detects a problem, as in the previous section V-D, then as described in (III-E), the next step is to analyze the situation to determine what the nature and source of the problem are. While the **rpm** can potentially provide a better sensorimotor simulator, which might be the essential tool for carrying out such an analysis, I don't know what it is about the nature of the **rpm** that would suggest that it should have the cognitive capability to carry out such an analysis, unless it might be that there is something more complicated in the creation or maintenance of the models than appears at first. Either way, we come right up against a need for some cognitive analytical ability, which we could also loosely refer to as intelligence.

Deus in machina: Now the **rpm** has sufficient cognitive ability to assess the source and nature of the problem, and given that information, to determine a resolution to the situation. Don't get confused. The **rpm** addresses these issues by going through a cognitive reasoning process, using the simulator, etc., it just received its ability to do cognitive reasoning by magic.

VI. Related work

Works that we would consider most directly related to the ideas in this paper are primarily those that address creating and/or maintaining internal models in a sufficiently proprioceptive way as to allow the possibility of anomaly detection, regardless of whether that was the intended purpose. For example, the main piece of work in [25] is an architecture that simulates the architecture of multiple forward and inverse pairs from [46] [47]. Although all four of [12] [13] [14] [15] reduce to essentially the same algorithm for the hard part of what they have to do, they vary in the degree to which they do or don't get speed-up for using simulation, for example, or focus on discovering self-structure for the first time versus allowing a possible change of structure, as if from injury. These papers are probably the most closely connected to our work.

Finally, [43] is very interesting, because it starts with a subsumption system, and adds an adaptive sort of feature intended to monitor abnormal behavior, and ends up with a behavior based system. That makes it sort of parallel to what we do in this paper, in a "lite" version.

VII. Conclusions

At the start of the paper we wanted to know what an autonomous, behavior-based robot might have to gain from going along with the idea in [23] to log all its sensory, state, and behavior data. Now we know: A very good shot at anomaly detection; with a good chance of becoming much more perturbation tolerant over time *if* the modeler can dig up some cognition somewhere. The related work suggests that if we start with the perspective of, say, wanting to produce a prototype architecture as proof of concept, it would probably be very worthwhile to do.

Another question

At the end of section V, in subsection V-E, we had an autonomous behavior-based robot combined with an intelligent proprioception modeler of same, or call it cognitively enhanced if you prefer. At first glance it seemed a bit like a classic hybrid, there are some superficial resemblances, but in this case the non-behavior-based component knows nothing about planning, it knows only about the body and behavior of the behavior-based component—which it fixes and adjusts to make the latter more robust and perturbation-tolerant. So there doesn't seem to be so much of a "Robotic Body-Mind Integration Problem" either [28]. It might be interesting to see how this compares with other architectures, *cf.* [17], but we will have to leave it for another time; perhaps after it has moved from the cloudy world of *gedankenexperimenten* into an embodied and situated form.

Acknowledgments

As I mentioned in the introduction, the question that prompted the investigation that led to this paper was sparked by D. Gage, whom I thank also for his interest and some interesting conversations. One of the most important influences on how I thought about and developed the problem was a conversation with A. Sloman at AAAI-06 about a month *before* Gage spoke at PerMIS'06; I thank Sloman most heartily for that. I also thank M.L. Anderson for some quite relevant and useful conversation. My sincere appreciation to P. Coronado and NASA as well as R. Madhavan and the Workshop Organizers for their kind flexibility as the work on this paper was being completed. And finally, deeply heartfelt thanks to E. Atkins for unrealized-at-the-time influence in starting me on this path, and to A. Agrawala for helping me continue.

REFERENCES

[1] Albus, J.S., "A New Approach to Manipulator Control: the Cerebellar Model Articulation Controller (CMAC)," *Trans. ASME, Series G. Journal of Dynamic Systems, Measurement and Control* **97** (1975), 220–233.

[2] Anderson, M.L., "Embodied cognition: a field guide." *Artificial Intelligence* **149** (2003), 91–130.

[3] Anderson, M.L., "Evolution of cognitive function via redeployment of brain areas," *The Neuroscientist* **13** (2007), 13–21.

[4] Anderson, M.L., Perlis, D.R., "Logic, self-awareness and self-improvement: the metacognitive loop and the problem of brittleness," *Journal of Logic and Computation* **15** (2005), 21–40.

[5] Anderson, M.L., Oates, T., Chong, W., Perlis, D, "The metacogntive loop I: enhancing reinforcement learning with metacognitive monitoring and control for improved perturbation tolerance," *Journal of Experimental and Theoretical Artificial intelligence* **18** (2006), 387–411.

[6] Arkin, R.C., *Behavior-Based Robotics.* Cambridge, MA, London: The MIT Press, 1998.

[7] Bays, P.M., Wolpert, D.M., "Actions and consequences in bimanual interaction are represented in different coordinate systems," *The Journal of Neuroscience* **26** (2006), 7121–7126.

[8] Berthoz, A., *The Brain's Sense of Movement.* Translated by G. Weiss. Cambridge, MA, London: Harvard University Press, 2000.

[9] Bhushan, N., Shadmehr, R., "Computational nature of human adaptive control during learning of reaching movements in force fields," *Biological Cybernetics* **81** (1999), 39–60.

[10] Bizzi, E., Mussa-Ivaldi, F.A., "Neural basis of motor control and its cognitive implications," *Trends in Cognitive Sciences* **2** (1998), 97–102.

[11] Blakemore, S.J., Goodbody, S.J., Wolpert, D.M., "Predicting the consequences of our own actions: the role of sensorimotor context estimation," *The Journal of Neuroscience* **18** (1998), 7511–7518.

[12] Bongard, J.C., Zykov, V., Lipson, H., "Automated synthesis of body schema using multiple sensor modalities," in Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems, Bloomington, IN, USA (ALIFEX), 2006, pp. 220–226.

[13] Bongard, J.C., Zykov, V., Lipson, H., "Resilient machines through continuous self-modeling," *Science* **314** (2006), 1118–1121.

[14] Bongard, J.C., Lipson, H., "Automated damage diagnosis and recovery for remote robotics," in Proceedings of the 2004 International Conference on Robotics and Automation (ICRA), Omnipress, pp. 3545–3550.

[15] Bongard, J.C., Lipson, H., "Automatic synthesis of multiple internal models through active exploration," in AAAI Fall Symposium: From Reactive to Anticipatory Cognitive Embodied Systems, 2005.

[16] Brooks, R.A., *Cambrian intelligence: The Early History of the New AI.* Cambridge, MA, London: MIT Press, A Bradford Book,1999.

[17] Bryson, J., "Cross-paradigm analysis of autonomous agent architecture," *Journal of Experimental and Theoretical Artificial intelligence* **12** (2000), 165–189.

[18] Bugmann, G., "Role of the cerebellum in time-critical goal-oriented behaviour: anatomical basis and control principle," in *Emergent Neural Computational Architectures Based on Neuroscience: Towards Neuroscience-inspired Computing,* (LNAI 2036) S. Wermter, J. Austin, D. Willshaw, eds. Berlin, Heidelberg, New York: Springer, 2001, pp. 255–269.

[19] Clancy, W.J., *Situated Cognition: On Human Knowledge and Computer Representations.* Cambridge, UK, New York: Cambridge University Press, 1997.

[20] Cole, J., Paillard, J., "Living without touch and peripheral information about body position and movement: studies with deafferented subjects," in *The Body and the Self,* J.L. Bermúdez, A. Marcel, N. Eilan, eds., Cambridge, MA, London: MIT Press, A Bradford Book, 1995, pp. 245–266.

[21] Davidson, P.R., Wolpert, D.M., "Widespread access to predictive models in the motor system: a short review," *Journal of Neural Engineering* **2** (2005), S313–S319.

[22] Elliott, T., Shadbolt, N.R., "Developmental robotics: manifesto and application," *Philosophical Transactions Royal Society London,* Part A **361** 2003, 2187–2206.

[23] Gage, D.W., "Meaninful metrics and evaluation of embodied, situated, and taskable systems," in Performance Standards for intelligent Systems Workshop (Gaithersburg, Maryland, USA), 2006, pp. 52–53.

[24] Gallagher, S., "Bodily self-awareness and object-perception," *Theoria et Historia Scientiarum: International Journal for Interdisciplinary Studies (Poland)* **7** (2003), 53–68.

[25] Haruno, M., Wolpert, D.M., Kawato, M., "MOSAIC model for sensorimotor learning and control," *Neural Computation* **13** (2001), 2201–2220.

[26] Holland, O., Goodman, R., "Robots with internal models: a route to machine consciousness?" *Journal of Consciousness Studies* **10** (2003) Special Issue: Machine Consciousness, pp. 2–45.

[27] Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., Kawato, M., "Modular organization of internal models of tools in the human cerebellum," *Proceedings of the National Academy of Sciences* **100** (2003), 5461–5466.

[28] Kawamura, K., Dodd, W., Ratanaswasd, P., "Robotic body-mind integration: next grand challenge in robotics," in Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication (Kurashiki, Okayama, Japan), 2004, pp. 23–28.

[29] Kawato, M., "Internal models for motor control and trajectory planning," *Current Opinion in Neurobiology* **9** (1999), 718–727.

[30] Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., Yoshioka, T., "Internal forward models in the cerebellum: fMRI study on grip force and load force coupling," *Progress in Brain Research* **142** (2003), 171–188.

[31] Kazennikov, O.V., Wiesendanger, M., "Goal synchronization of bimanual skills depends on proprioception," *Neuroscience Letters* **388** (2005), 153–156.

[32] Krakauer, J.W., Ghilardi, M.-F., Ghez, C., "Independent learning of internal models for kinematic and dynamic control of reaching," *Nature Neuroscience* **2** (1999), 1026–1031.

[33] Lacquaniti, F., Carrozzo, M., Ivanenko, Y.P., Zago, M., "On biological principles of motor control," in Proceedings of the First International IEEE EMBS Conference on Neural Engineering, Capri Island, Italy, 2003, pp. 454–457.

[34] Marr, D., "A theory of cerebellar cortex," *Journal of Physiology* **202** (1969), 437–470.

[35] Matarić, M.J., "Behavior-based robotics as a tool for synthesis of artificial behavior and analysis of natural behavior," *Trends in Cognitive Science* **2** (1998), 82–87.

[36] Miall, R.C., Weir, D.,J., Wolpert, D.M., Stein, J.F., "Is the cerebellum a Smith predictor?" *Journal of Motor Behavior* **25** (1993), 203–216.

[37] Minsky, M., "Matter, Mind and Models," `http://web.media.mit.edu/~minsky/papers/MatterMindModels.html`, last modified 1997.

[38] Naito, E., "Sensing limb movements in the motor cortex: how humans sense limb movement," *The Neuroscientist* **10** (2004), 73–82.

[39] Nowak, D.A., Glasauer, S., Hermsdörfer, J., "How predicitive is grip force control in the complete absence of somatosensory feedback?" *Brain* **127** (2004), 182–192.

[40] Peters, J., van der Smagt, P., "Searching a scalable approach to cerebellar based control," *Applied intelligence* **17** (2002), 11–33.

[41] Pfiefer, R., Scheier, C., *Understanding intelligence.* Cambridge, MA: MIT Press, 1999.

[42] Pipereit, K., Bock, O., Vercher, J.-L., "The contribution of proprioceptive feedback to sensorimotor adaptation," *Experimental Brain Research* **174** (2006), 45–52.

[43] Suksakulchai, S., Wilkes, D.M., Kawamura, K., "A behavior monitoring architecture for a mobile robot: enhancing intelligence to behavior-based robotics," in Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, 2003, pp. 157–162.

[44] Wikipedia, "Optic Nerve," `http://en.wikipedia.org/wiki/Optic_nerve`.

[45] Wolpert, D.M., Ghahramani, Z., Jordan, M.I., "An internal model for sensorimotor integration," *Science* **269** (1995), 1880–1882.

[46] Wolpert, D.M., Kawato, M., "Multiple paired forward and inverse models for motor control," *Neural Networks* **11** (1998), 1317–1329.

[47] Wolpert, D.M., Miall, R.C., Kawato, M., "Internal models in the cerebellum," *Trends in Cognitive Sciences* **2** (1998), 338–347.

[48] Ziemke, T., "Are robots embodied?" in Proceedings of the First International Workshop on Epigenetic Robotics, Lund, Sweden: Lund University Cognitive Studies **85**, 2001.