2004

Robot Self-Motivation: Balancing "Boredom" and "Confusion"

James Marshall

Doug Blank
Bryn Mawr College, dblank@brynmawr.edu

Lisa Meeden

Custom Citation

Marshall, J., Blank, D., and Meeden, L. (2004). Robot Self-Motivation: Balancing "Boredom" and "Confusion". A two-page poster of an idea. Unpublished.

Robot Self-Motivation: Balancing "Boredom" and "Confusion"

James B. Marshall
Computer Science
Pomona College
Claremont, CA 91711
marshall@cs.pomona.edu

Douglas Blank
Computer Science
Bryn Mawr College
Bryn Mawr, PA 19010
dblank@cs.brynmawr.edu

Lisa Meeden
Computer Science
Swarthmore College
Swarthmore, PA 19081
meeden@cs.swarthmore.edu

Abstract

This work focuses on the role of *self-motivation* in the developmental learning process of a mobile robot. We are interested in developing a general learning architecture that will enable a robot to build up hierarchical representations of its experiences through the processes of *abstraction*, for example by learning topological maps of sensory states, and *anticipation*, in which the robot learns to predict the outcome of applying its effectors to its current situation. Although abstraction and anticipation are active research areas (see, for example, (Kuipers and Beeson, 2001, Butz et al., 2002)), we believe that self-motivation is the missing component needed to create a successful epigenetic robotics system (Blank et al., 2002).

Furthermore, we believe that abstraction, anticipation, and self-motivation are inextricably intertwined and must develop together from the start. The learning of abstractions should be driven by the robot's attempts to anticipate the effects of its sensorimotor interactions with the environment, guided by its own internal motives. Our goal is to create a system that is capable of learning the characteristics not only of its surrounding environment, but also of its own sensors and effectors. This avoids the problem of anthropomorphic bias, since the system learns about its own capabilities through firsthand experience, rather than being provided with this knowledge by the designer at the outset. This also allows the system's knowledge of itself and its environment to be grounded in its own sensorimotor experiences.

The processes of abstraction and anticipation should be driven by motivations arising from within the system itself, rather than being imposed from the outside. Self-motivation can be viewed as an instance of the more general concept of *self-regulation*, loosely defined as arising whenever the behavior of a system depends in part on information or control signals originating from within the system itself, rather than from outside of it. In the context of robot learning, we use self-regulation as a mechanism through which a robot can learn to selectively vary its exposure to different parts of the environment (or to different regions of "mental space"). Our robot's self-regulation is based on an internally-generated measure of the error associated with a particular innate behavior. Based on its current sensor and motor readings, the robot learns to predict its next sensory state along with an estimate of the error contained in this prediction. This error estimate, generated by the robot itself, in turn serves as the basis for an internal or external effect. For example, an internal effect might be that of lowering the robot's learning rate when hard-to-predict situations are encountered. An external effect might be that of directly changing the orientation of the robot's wheels. The goal is for the robot to learn to recognize the regularities present in the environment while ignoring aspects of the environment that are inherently unpredictable. This effectively implements an attentional mechanism under the control of the robot itself.

We see self-motivation as a strong version of self-regulation in which the system is driven to continually push its mental development to create further structure and organization, while simultaneously pushing its physical self into new and unknown environments. For example, if the robot finds itself in a situation in which it cannot predict what will happen next, it may slow down or stop in order to allow itself to become more familiar with the situation. On the other hand, when prediction becomes very accurate, so that the robot is "bored", it may inject a degree of randomness or some other type of variation into its behavior in order to broaden its exploration of the environment. The net result is that when properly balanced, these tendencies can create an "edge of chaos" effect in which the robot seeks out situations and experiences that lie between the extremes of complete predictability and apparent chaos (Langton, 1990).

To explore these effects, we ran an experiment on a simulated Pioneer using the Stage simulator (Gerkey et al., 2003) and Pyro as our control layer

(Blank et al., 2003). Guided by a simple innate obstacle-avoiding behavior, the robot explored its environment while learning to predict future sensory states together with an estimate of their accuracy, using a feed-forward neural network. Furthermore, the robot continually monitored the actual error signals associated with these predictions and used this information to occasionally override its own innate behavior. The experiment defined three different error regimes. If the total error of the robot's prediction was sufficiently low, the robot's motivational state was considered to be "bored", since the low error indicated that the robot had successfully learned to anticipate the outcome of its current behavior. As a result, the robot moved randomly instead of following its "instincts", which often caused it to experience new, unfamiliar situations which were not yet predictable. On the other hand, if the error was sufficiently high, the robot was considered to be "overwhelmed". In this case, its response was to simply stop moving. This gave it time to learn to at least predict its current (static) situation correctly before resuming exploration. Otherwise, the robot was in a "balanced" state that was not overly chaotic but not entirely predictable. This latter region is where we believe the most productive learning can occur.

Figure 1 shows a representative graph of the motivational categories experienced by the robot over time as it explored its environment. Each bar shows the relative percentage of time the robot spent in each state, measured over a fixed time period. At the beginning of the run, the robot displays tentative and halting behavior, since much of its time is spent in the "overwhelmed" state, simply trying to learn what to expect when at rest. Gradually, this gives way to greater movement and more confident-seeming behavior, as the proportion of balanced and "bored" states increases. By the end of the run, the robot is rarely unable to anticipate what will happen next, and often overrides its innate behavior in favor of random exploration.

Although the work described here is ongoing and preliminary, it may offer a hint of the potential role that self-motivation can play in developmental robotics, and possibly other more mundane tasks as well. Many avenues remain to be explored, such as using different types of internally-generated feedback signals as the basis for self-control, or varying the ways in which this feedback signal can affect the behavior of the system. For instance, we are investigating the effects of having the robot vary its learning rate in response to its current motivational state, although the results so far have been inconclusive. Another interesting possibility would be to incorporate temporal information into the robot's predictive mechanisms by using a recurrent neural network in place of the feed-forward network.

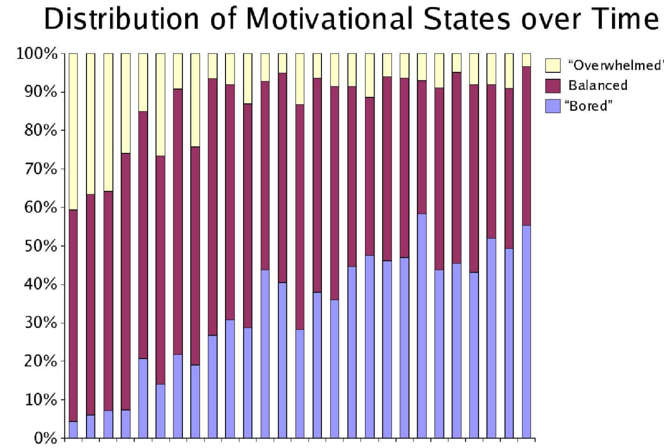

Figure 1: The relative distribution of motivational states over 13,500 steps. At the beginning of the run, the "overwhelmed" category (on top) accounts for about a 40% share of the states, but slowly subsides. The "bored" category begins at less than 5%, but steadily grows over the run. The percentage of balanced states remains fairly stable over the run.

References

Blank, D., Meeden, L., and Kumar, D. (2002). Bringing up robot: Fundamental mechanisms for creating a self-motivating, self-organizing architecture. In *Workshop proceedings of* Growing up artifacts that live*, at Simulation of Adaptive Behavior*.

Blank, D., Meeden, L., and Kumar, D. (2003). Python robotics: An environment for exploring robotics beyond LEGOs. In *ACM Special Interest Group: Computer Science Education Conference (SIGCSE)*.

Butz, M. V., Sigaud, O., and Gerard, P. (2002). Internal models and anticipations in adaptive learning systems. In *Proceedings of the Workshop on Adaptive Behavior in Anticipatory Learning Systems*, pages 1–23.

Gerkey, B., Vaughan, R., and Howard, A. (2003). The Player/Stage project: Tools for multi-robot and distributed sensor systems. In *Proceedings of the International Conference on Advanced Robotics*, pages 317–323.

Kuipers, B. and Beeson, P. (2001). Toward bootstrap learning for place recognition. In *Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems: Papers from the 2001 AAAI Fall Symposium*, pages 25–30. AAAI Press.

Langton, C. G. (1990). Computation at the edge of chaos. *Physica D*, 42:12–37.