DOI: http://dx.doi.org/10.7551/978-0-262-32621-6-ch151

Pleasure, Persistence and Opportunism in Action Selection

Matthew Lewis[1], Antoine Hiolle[1] and Lola Cañamero[1]

[1]Embodied Emotion, Cognition and (Inter-)Action Lab
School of Computer Science, University of Hertfordshire
College Lane, Hatfield, Herts, AL10 9AB, U.K.
matt-l@semiprime.com, L.Canamero@herts.ac.uk

Abstract

An autonomous robot must show appropriate levels of persistence and opportunism to survive. We address this problem by using a mechanism akin to pleasure that modulates exteroception as a function of need satisfaction, rather than based on internal deficits and external threats as in previous work. The different context in which the modulating hormone is released has important consequences on persistence and opportunism.

Introduction

Persistence and opportunism are two key features of action selection architectures (Tyrrell, 1993; Maes, 1995). For an autonomous robot that has to satisfy multiple conflicting survival-related needs, it is crucial not only to choose behaviors that do so in a timely fashion, but also to *persist* in their execution for long enough to guarantee sufficient satisfaction. Persistence is important to avoid what is known as the "dithering" problem, which occurs when a robot keeps switching between trying to satisfy two needs without satisfying either of them enough to guarantee survival. Another key feature is opportunism: to consume a resource that might not be needed at present but is available now and might not be available later. The degree to which a robot should show persistence and opportunism depends on multiple factors; we could generally say that persistence leads to a more "conservative" action selection behavior and opportunism to a more "risky" one. In previous work (Avila-García and Cañamero, 2004), we showed that persistence and opportunism can also become negative when done in excess, and proposed a mechanism inspired by emotions in natural systems and based on hormonal modulation of the perception of external stimuli (the resources), to address these problems. In that work, the hormone modulating perception was released as a function of internal deficits and the presence of threats in the environment, i.e., it was released signaling that *things were not functioning well*. Building on that work, here we present a related motivated action selection architecture that uses a mechanism akin to pleasure (also modeled using artificial hormones) to modulate the perception of the resources. Following the principle that pleasure *signals well functioning* (Panksepp, 1998), the "pleasure hormone" is released as a function of need satisfaction. The very different context in which the hormone that modulates perception is released creates different behavioral dynamics and has important consequences regarding how the architecture addresses persistence and opportunism.

Robot Architecture

Following a definition of autonomy as self-regulation, we use a homeostatically-controlled motivated architecture having to solve a two-resource action selection problem, in the same vein as Avila-García and Cañamero (2004). Our robot (an Aldebaran humanoid Nao) has thus two homeostatic internal variables – energy and moisture deficits, ranging between 0 and 100, where 100 is the fatal limit – that, in conjunction with the stimuli in the environment, motivate it to select the appropriate behavior to execute. Perception of the external environment is done using a camera to detect the resources – colored balls used as "food" and "drink" – and sonar and contact sensors to move around safely. Each resource, when consumed by the robot, reduces the corresponding simulated deficit. Fig. 1 shows the experimental setup and Fig. 2 the architecture. Two motivations – *thirst* and *hunger* – each corresponding to a deficit, combine the perception of internal deficits and external stimuli to guide behavior selection. As in Avila-García and Cañamero (2004) we calculate each motivation level using the formula

$$motivation_i = deficit_i + (deficit_i \times \alpha \times cue_i) \qquad (1)$$

where cue_i is the magnitude of the stimulus – here perceived size of a resource related to the deficit – and α modulates the perception of resources. In this case, we have used a "pleasure" hormone that we hypothesized can allow persistence while reducing pathological opportunistic behavior. Motivations influence which behaviors are selected for execution (in 8Hz-tick cycles) by passing them their activation level. Each behavior – a perception-action loop – has a self-controlled activation level that depends on sensory inputs that are relevant for it, and activation threshold, and becomes potentially executable when its activation level is above the threshold. For this study we have implemented two top-level behavioral systems – "behaviors" for short – each linked

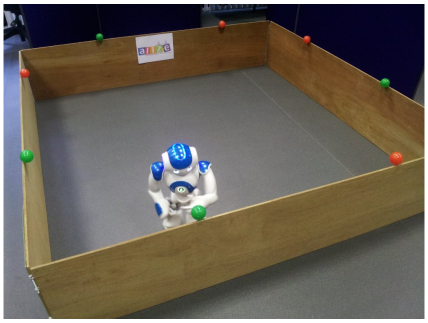

Figure 1: Experimental setup. The green balls around the arena represent drink resources, and the red food resources.

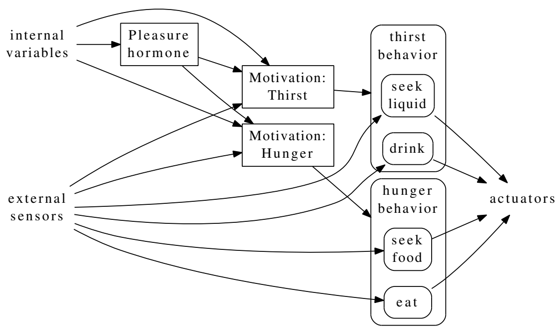

Figure 2: Simplified robot architecture

to a motivation. These *thirst* and *hunger* top-level behaviors are implemented by embedding simpler behaviors, both consummatory (*eat* food and *drink*, which can be executed only if the corresponding stimulus is present) and appetitive (goal-seeking, including wandering randomly, scanning the environment, tracking a detected object, and walking in the direction that the head is turned), to seek out the relevant environmental resources, and consume them if found. The selected top-level behavior in turn acts as a behavior selection mechanism selecting from amongst its sub-behaviors. While the top-level behaviors are selected on a winner-take-all basis by motivations, multiple lower-level behaviors can be run simultaneously, provided they use disjoint sets of actuators.

Experiments and Results

To investigate our hypothesis, we modulated exteroception using a "pleasure" hormone released when either of our deficits decreases. Subsequently, the level of the hormone decays back to its basal level. Intuitively, the hormone indicates that recent behaviors were beneficial. We tested two conditions: a modulated α in equation 1 proportional to the level of the hormone, and a control condition using a constant α. Results from two runs are shown in Fig. 3. In the case where α is *constant* (left), the robot displays adaptive persistence in its drinking behavior soon after starting (Food deficit in the 40s). However, not long after this, with the Food deficit around 60 it eats, and keeps eating opportunistically while the Drink Deficit continues to rise. The robot

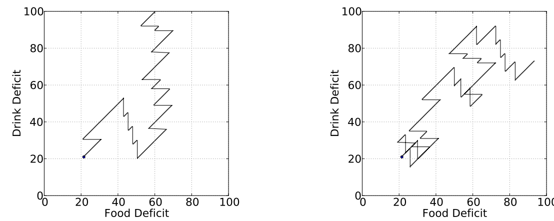

Figure 3: Results of two runs showing the evolution of deficits over time. Left: constant $\alpha = 0.5$ (control). Right: α modulated by pleasure hormone.

dies of thirst in less than four minutes. In the case where α is *modulated* by the pleasure hormone (right), the robot shows persistence (particularly to the top right of the graph), but avoids getting stuck performing behaviors opportunistically. Although it has high Food and Drink deficits, it is still alive after more than five minutes.

Discussion

Our preliminary results are consistent with the hypothesis that in some environments (excessive) opportunistic behavior can be maladaptive, and that using pleasure to modulate the α parameter in our motivation formula can offer a solution to this problem, while still allowing persistence. Using pleasure in this way means that opportunistic behavior can be reduced overall, but cues that are present during the release of the hormone will have their significance magnified until the hormone has decayed back to its basal level. Since it is likely (but not guaranteed) that these cues influenced the behaviors that caused the drop in the deficits, and the subsequent release of the hormone, pleasure would be adaptively communicating the instruction "keep acting in response to any cues that are present". This is consistent with the role of positive affect in fostering openness towards the world. This also opens up the possibility of using other sources of pleasure to make behaviors persist, i.e. to externally influence its actions. When interacting with a human, social pleasure could be used as a tool to reinforce behaviors in the robot.

Acknowledgments

This work was supported by EC grant FP7-ICT-248116 (ALIZ-E). We are grateful to the other members of the Embodied Emotion, Cognition and (Inter-)Action Lab for discussions on this research. The opinions expressed are solely the authors'.

References

Avila-García, O. and Cañamero, L. (2004). Using hormonal feedback to modulate action selection in a competitive scenario. In *Proc. SAB'04*, pages 243–252. MIT Press.

Maes, P. (1995). Modeling adaptive autonomous agents. *Artificial Life: An Overview*, pages 135–162.

Panksepp, J. (1998). *Affective Neuroscience*. OUP.

Tyrrell, T. (1993). *Computational Mechanisms for Action Selection*. PhD thesis, University of Edinburgh.