Developing Affect-Modulated Behaviors: Stability, Exploration, Exploitation or Imitation?

Arnaud J. Blanchard   **Lola Cañamero**
Adaptive System Research Group
School of Computer Science
University of Hertfordshire
College Lane, Hatfield, Herts AL10 9AB
{A.J.Blanchard, L.Canamero}@herts.ac.uk

Abstract

Exploring the environment is essential for autonomous agents to learn new things and to consolidate past experiences and apply them to improve behavior. However, exploration is also risky as it exposes the agent to unknown, potentially overwhelming or dangerous situations. A trade-off must hence exist between activities such as seeking stability, autonomous exploration of the environment, imitation of novel actions performed by another agens, and taking advantage of opportunities offered by new situations and events. In this paper, we present a Perception-Action robotic architecture that achieves this trade-off on the grounds of modulatory mechanisms based on notions of "well-being" and "affect". We have implemented and tested this architecture using a Koala robot, and we present and discuss behavior of the robot in different contexts.

1. Introduction

In previous work (Blanchard and Cañamero, 2005a) we proposed a robotic architecture to develop affective bonds to a caretaker. However, in that architecture the robot was not able to discover any new situation in the absence of external stimulation, as it only tried to stay safe with the caretaker. For autonomous agents (children, animals or robots), exploring the environment is essential as it allows them to learn new things and also offers opportunities to consolidate past experiences and apply them to improve their behavior in situations similar to other situations previously experienced. Exploration is also a risky activity as it exposes the agent to unknown, potentially overwhelming or dangerous situations. Therefore, exploration must be done with care, and a trade-off should exist between activities such as seeking stability, exploring, imitating novel actions performed by another agent, or trying to take advantage of opportunities offered by new situations and events. Endowing a robot with this capability poses three main problems: (1) generating all these behaviors from the same underlying architecture; (2) autonomously switching among them; and (3) achieving a good balance in the execution of these activities. We present a robotic architecture based on biologically plausible principles to achieve these three goals. In addition, trying to solve the exploration, exploitation problem leads to low-level imitation as aside-effect in our architecture.

This architecture, which we have implemented and tested using a Koala robot (http://k-team.com/robots/koala), follows a "Perception-Action" (PerAc) approach rooted both in psychology (Prinz, 1997) and in robotics (Gaussier and Zrehen, 1995), that we have already used to model various phenomena in previous work (Blanchard and Cañamero, 2005a, 2005b, 2006). The PerAc approach postulates that perception and action are tightly coupled and encoded at the same level. Action is thus executed as a "side-effect" of wanting to achieve, improve or correct some perceptions. The perception-action loop can be seen in terms of homeostatic control (Ashby, 1952), according to which behavior is executed to correct perceptual errors. Actions that allow to correct different perceptual errors are selected on the grounds of sensorimotor associations[1]. It is important to notice that the terms sensations and perceptions are often used interchangeably in the literature. In this paper we use *sensation* to denote the raw input from the sensors and *perception* to denote the interpretation (including associated actions or affordances) of the sensation.

Taking an incremental approach to design our architecture, we first recall (section 2.) the architec-

---

[1]Sensorimotor associations can be "hardcoded" by the designer (e.g., with static coefficients, as it is our case here) or learned from experience by the robot—see e.g., (Gaussier et al., 1998, Andry et al., 2003) for examples applied to imitation of movements by a robot.

ture presented in (Blanchard and Cañamero, 2005a) adopted here to seek stability to later add elements that progressively induce capabilities of exploration, exploitation, and finally low-level imitation. At each step, new behavioral capabilities are added while preserving the existing ones. The robot must thus achieve a good balance among all its activities.

Autonomously achieving an adaptive execution of activities can be seen as an instance of the behavior selection problem. In our case, however, changes in observable behavior are not achieved by "switching" among a set of discrete behaviors, but differential execution of activities relies on a modulatory mechanism based on notions of "well-being" and "affect." In our model, well-being depends on the internal (physiological) states of the agent, whereas affect depends on the values of its sensors (sensations). Well-being thus corresponds to the endogenous factor of comfort described by Dunn (Dunn, 1977), whereas affect corresponds to its exogenous factor. We use the following definitions:

- The *well-being* of an agent is a measure (taking values between 0 and 1) of the viability of its internal state, i.e., the distance of the variables composing the internal state of the agent to their ideal values. A high level of well-being corresponds to a zone of good viability in its physiological space, as depicted in Figure 1 (left).

- *Affect* in this model is the evaluation (expressed in values between 0 and 1) of "goodness" or "safety" of a situation based on the familiarity (in terms of frequency) and the past well-being (pleasantness) of the associated sensation. A high level of affect corresponds to the fact that the agent evaluates the situation in the world as highly safe; as depicted in Figure 1 (right), this can be represented in the agent's sensory space as a function of the mismatch between the actual sensation and a past "ideal sensation" (called "desired sensation" in our model, see below) to which the robot would tend to return. Using affect, agents are able to evaluate how far they are from a safety zone which corresponds to a familiar zone or to a zone where they expect to maximize their well-being and therefore their life time (Likhachev and Arkin, 2000).

As we will show in the remainder of the paper, modulation of the architecture based on well-being and affect achieves adapted execution of different activities. This makes the robot explore if nothing happens, take advantage of opportunities in the presence of novelty while avoiding danger, imitate another agent that is performing a novel action, or return to familiar situations ("safe zones") if nothing happens but the well-being is low.
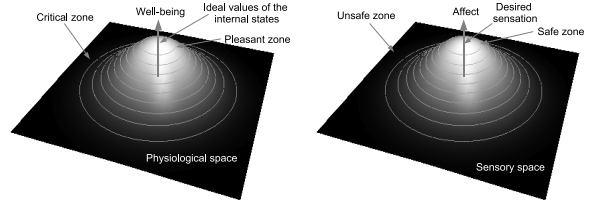

Figure 1: The two components of the robot's comfort: well-being can be represented in the physiological space (left) and affect in the sensory space (right).

2. Seeking Stability

In early life, animals are often placed in a safe environment (e.g., a nest), in which a caretaker (usually the mother) looks after them; therefor good behavior would be to try to remain in or to return to this situation. The phenomenon called "imprinting" that Konrad Lorenz observed in geese in the 1930's, and by which some animals become attached to the first object they perceive, achieves this effect. However, the initial situation might not always be ideal, or it might change due to external events, changes in interaction with the caretaker, etc.; therefore, it is not always possible or beneficial to try to maintain or return to the initial situation, and adaptation to the new circumstances becomes necessary. In (Blanchard and Cañamero, 2005a), we proposed an architecture that combined imprinting and adaptation, following (Bateson, 2000), and that we have adapted here to model stability-seeking behavior.

In our setup (Figure 2), we use a Koala robot which we place at the front of the caretaker. The only information that the robot has about the objects around it, including the caretaker, is the sensation of distance (Sd), obtained via its infrared (IR) sensors. Stimulation of one its lateral IR sensors increases the internal well-being (Wb) of the robot (which decreases in the absence of stimulation), and the robot will execute behaviors that allow to keep the value of this internal (physiological) variable close to its ideal value.
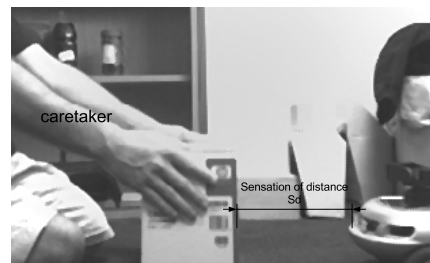

Figure 2: Setup of the experiments.

To learn which sensation corresponds to a safe zone of the robot (what we call here

"desired sensation"), we first proposed in (Blanchard and Cañamero, 2005a) that this sensation should be the most familiar one. To compute the most familiar sensation of distance ($\overline{Sd}$), we compute the average sensations the robot had in its life using the simple incremental learning rule (1) similar to the one used to model conditioning (Rescorla and Wagner, 1972), where t represents the time from the start and $\eta(t) = \frac{1}{t}$, the learning rate:

$$\overline{Sd(t)} = \overline{Sd(t-1)} + \eta(t) \times (Sd(t) - \overline{Sd(t-1)}) \quad (1)$$

However, the robot should not give the same importance to all the sensations it had but should give more importance to the sensations which were associated to higher well-being. To implement this, we redefine the learning rate (η) as the actual well-being (Wb) divided by the sum ($\tilde{W}b$) of all the values of its well-being in the past. Learning is not dependent on the absolute value of well-being but on its variation, which means that when the well-being is constant, regardless of its value, we have exactly the same learning rule used to learn the average sensation, but when the well-being changes, more or less weight is given to the final desired sensation. The problem is that, with the time, $\tilde{W}b$ will become very large, and the robot will learn very slowly and even not learn at all. To give more weight to recent sensations, we propose to associate different desired sensations (Sd_k) with different time scales (k) using different learning rate (η_k) where $\eta(t)_k = \left(\frac{Wb(t)}{\tilde{W}b(t)}\right)^k$. Porr (Porr and Worgotter, 2003) also uses different time scales for learning, but in our case we do not try to learn in order to predict sequences and the well-being (or reward) has an influence on the learning rate. With long-term time scales (k tends to $+\infty$), the learning rate is very slow, and the first sensations have a strong effect on the desired sensation of this scale. With a short-term time scale (k tends to 0), the learning rate is high and strongly dependent on the well-being, the actual sensation has a strong effect on the desired sensation of this scale. We have all the intermediate effects with the intermediate time scales and in all our experiments, we use ten time scales where k varies from 0.2 to 2 by steps of 0.2.

The actions corresponding to the difference between the actual sensation and the desired sensations can be perceived either using an association learned during a babbling process (Andry et al., 2003) or using an hardcoded association as we do. In this hardcoded association, the perception of the action which could have produced the difference between the actual and the desired sensation of distance to the caretaker corresponds to the action (speed of movement) of moving forward or backward depending on the sign and amplitude of the difference. However, as we use different time scales, the same difference in
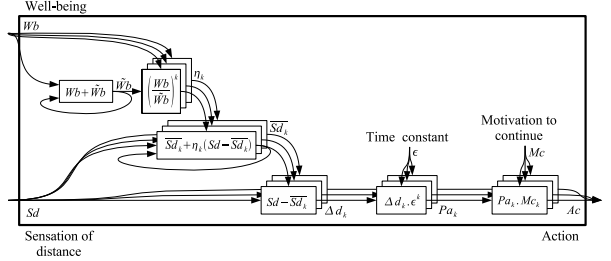

Figure 3: Architecture to generate stability-seeking behavior.

sensation can correspond to a fast action in a short-term time scale and to a slow action in a long-term time scale. Therefore, to compute the perceived actions, we modulate the difference (Δd_k) between the actual and the desired sensation with a time constant (ϵ) to the power of k, i.e. adapted to each time scale. The perceived action (Pa_k) at each time scale is defined:

$$Pa_k(t) = \epsilon^k \times \Delta d_k(t) \quad (2)$$

where $\Delta d_k(t) = Sd(t) - \overline{Sd_k(t)}$. In order to try to reach the desired sensation, the robot should oppose itself to these perceived actions, and its motivation to continue (Mc_k) the perceived actions must be negative (we use -1). As shown in Figure 3, the final action (Ac_k) result from the average perceived actions (Pa_k) modulated by the motivation to continue (Mc_k), and in this case the robot always refuses new situations and keeps stability:

$$Ac_k(t) = Pa_k(t) \times Mc_k(t) \quad (3)$$

3. Exploration

Reproducing the imprinting phenomenon and adding the possibility of adaptation allows the robot to memorize which are the sensations it should try to reach in order to have good stability and well-being. However, in the absence of external stimulation, the robot described in (Blanchard and Cañamero, 2005a) will stay all the time in the same situation as it will never have the possibility to experiment the effects of different sensations. On the contrary, animals often look for novelty (Panksepp, 1999, Power, 2000) and, as already pointed out by (Oudeyer and Kaplan, 2004, Kaplan and Oudeyer, 2004, Steels, 2004), it would be very beneficial for robots to look for novelty—in our case, unfamiliar sensations. Obtaining new sensations can however be dangerous and not always useful as too much novelty does not produce efficient learning (Kaplan and Oudeyer, 2005). Moreover, Dunn (Dunn, 1977) observed that children explore more when they are in a familiar environment;

Likhachev and Arkin (Likhachev and Arkin, 2000) use the zone of comfort in order to modulate exploration in robots.

To generate spontaneous exploration, in (Blanchard and Cañamero, 2006) we proposed to increase the effect of an exploratory behavior while the robot does not have any other specific motivation. Exploratory behavior can consist in the execution of different actions selected randomly (Andry et al., 2003), or selected in order to lead to a maximum of learning (Kaplan and Oudeyer, 2004), or as in our case simply the action (Be) of moving forward where Be represents the innate speed of the movement for exploration. In this paper we use the variable *apathy* (Ap) to denote a state in which the robot does not have any specific motivation, and therefore the motivation to explore (Ae) is initiated by the exploratory behavior (Be) and increases while apathy is high.

$$Ae(t) = (Ae(t-1) + Be(t)) \times Ap(t) \qquad (4)$$

For different reasons that we will see later, the robot can have the motivation to stop, cancel or amplify the ongoing actions. We use a variable Mc to denote the motivation to continue. Mc is negative when the robot tries to oppose itself to the ongoing actions, and positive when it tries to amplify them. The apathy (Ap) leading to exploration represents the case in which the robot has low motivation to act (to continue or to stop):

$$Ap(t) = e^{-r \times (Mc(t))^2} \qquad (5)$$

where r is a parameter defining the decay rate of apathy as a function of motivation.

The robot can now theoretically explore (move forward in order to experiment non-desired or novel sensations) when it does not have any other motivation. However, the behavior described in the section 2., always carries the motivation to avoid new sensations (motivation to continue constant and equal to -1), and therefore the robot would always be opposed to the exploratory behavior. Our proposed solution to this problem is to apply the notion of *affect* presented earlier to dynamically modify the motivation to continue. Affect reflects the proximity (in the sensory space) of the robot to its desired sensation, but the importance of a distance to a particular sensation of the robot is correlated with the actions that the robot has to do in order to reach that sensation. This measure of distance is given by Pa_k, representing the perceived actions needed to make the robot reach its desired sensation as. The variable affect (Af) varies from 1 (close to the desired sensation) to 0 (far to the desired sensation) and is defined in:

$$Af_k(t) = e^{-s \times (Pa_k(t))^2} \qquad (6)$$

where s is a parameter indicating the decay rate of the affect as a function of the perceived actions needed to reach the desired sensation.

We recall that when the affect is high, the robot can (and must, to perform exploration) afford to increase the effects of perceived actions, whereas it should try to decrease them when the affect is low. The motivation to continue must then be positive when the affect is high, and negative when the affect is low. The threshold between what should be considered as "low" or "high" affect is subjective and we set this threshold using a static parameter q defining the characteristic behavioral profile (or in a very restricted sense the "personality") of the robot and q can be interpreted as the *timorousness*. If q is high, the robot will often oppose itself to the perceived actions whereas if q is low, the robot will more often try to continue the perceived actions. This helps to define the degree of openness to the world (Op_k) which is a variable for each time scale representing the amplification of any perceived actions for "curiosity":

$$Op_k(t) = Af_k(t) - q \qquad (7)$$

In this part of the architecture, we equate the motivation to continue to the openness to the world; the architecture is similar to the previous one but the motivation to continue is computed as in Figure 4 instead of being equal to -1.
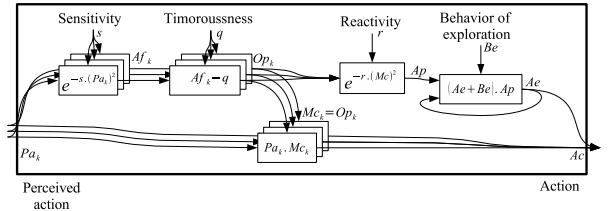

Figure 4: Generation of exploratory behavior.

When we apply this architecture to a robot in a static environment, it does not move at the beginning but when the motivation to explore (Ae) increases, it starts to move forward; in this case, the robot is not in a familiar situation anymore, and therefore it starts to have the motivation to oppose itself to the ongoing actions, the level of apathy (Ap) will become low and this will stop the exploration, and eventually (depending on r and s) cancel the last movement. It is interesting to note that this produces a behavior to advance by small steps; such "cautious approach" behavior is not only useful to control the level of unfamiliarity during exploration, but it also reproduces the approach of an animal to a new stimulus—it moves forward, stops, waits a bit, moves forward again and so on. However, if the caretaker moves, the robot will inhibit its exploratory behavior and only try to reach stability, i.e., retrieving its initial

distance to the caretaker. Figure 5 shows typical positions of our Koala robot during this "cautious approach" exploratory behavior when tested with three different values of q 0, 0.5 and 1, s set to 0.1, and r to 1.
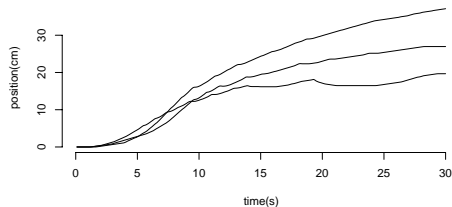

Figure 5: Successive positions of the robot during exploration for three different values of q—0, 0.5 and 1 for curves from top to bottom.

We can observe that the robot moves confidently (smoothly) when q is low, whereas it moves with hesitation (with some backwards movements) when q is high. In all cases, exploration was slower when the robot was farther away from its initial (and therefore familiar) situation (initial distance to the caretaker). The decay rate of exploration as a function of unfamiliarity depends on the parameter s and therefore can be changed.

4. Exploitation and Interruption-Related Behaviors

We have shown that the robot can autonomously modulate two kinds of behaviors—seeking stability and exploration—depending on external events. The first behavior makes the robot try to reach sensations know as familiar or pleasant, the second makes it explore new sensations when it is already in a familiar or desired environment. However, if during an action (executed either in order to reach a desired sensation or in order to explore), the well-being increases or decreases suddenly, the robot should interrupt the current behavior and increase or respectively decrease the effect of the perceived ongoing actions. In fact, if the robot accidentally moves close to a reward (stimulus increasing the well-being), it should continue its movement, showing *opportunism*. On the contrary, it should cancel or oppose itself to this movement if it discovers a danger (stimulus decreasing the well-being), therefore showing *avoidance* behavior. Affect is able to interrupt ongoing behavior like (Simon, 1967) observed it could be done in animals or human using a motivational and emotional control. The difficulty for the robot is to know if the variation of well-being is due to the recent changes or to long-term changes. Therefore, we use the average past well-being ($\overline{Wb_k}$) at different time scales to compare it with the actual well-being (Wb). In

fact, we use the same learning rates as those used for the computation of desired sensations in order to estimate the well-being associated to each desired sensation:

$$\overline{Wb_k(t)} = \overline{Wb_k(t-1)} + \eta^k \times (Wb(t) - \overline{Wb_k(t-1)})$$
(8)

We call *pleasure* (Pl_k) the measure of the variation (between -1 and 1) of well-being at different time scales:

$$Pl_k(t) = Wb(t) - \overline{Wb_k(t)}$$
(9)

We then use this notion of pleasure to compute the motivation to continue (Mc_k), which now not only depends on the openness to the world (Op_k) but also to the variation of well-being:

$$Mc_k(t) = Op_k(t) + Pl_k(t)$$
(10)

The new way of computing the motivation to continue generating opportunism and avoidance behaviors is presented in Figure 6.
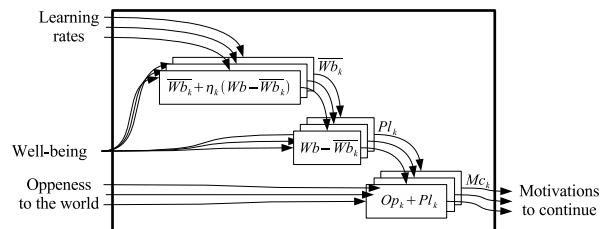

Figure 6: Computation of motivation to continue opportunism and avoidance behaviors.

To test the ability of the robot to take advantage of opportunities, we use the same experimental setup as in previous sections but this time we put boxes on the side of the path of the robot as rewards, as shown in Figure 7: when the robot moves forward, the boxes touching the contact sensor on the side of the robot increase the well-being. The values of the static parameters are similar to the previous experiment but s is set to 0.001 (a small s decreases the influence of the initial situation) and q to 0.75 (a large q decreases the influence of the exploratory behavior). We run a dozen of experiments in three different contexts. In the first experiment (exp 1) were there is no external reward—the well-being is constant, like in section 3.—only s and q have changed. In the second experiment (exp 2), there is a reward at the front of the robot. In the third experiment (exp 3), the robot is already next to a reward and by exploring will lose it.

Figure 8 presents the successive positions and well-being of the robot during one typical trial in each context (exp 1 in solid line, exp 2 in dashed line and exp 3 in dotted line). We observe that the robot always starts with an exploratory behavior similar to that in section 3.; however, if it meets a reward it
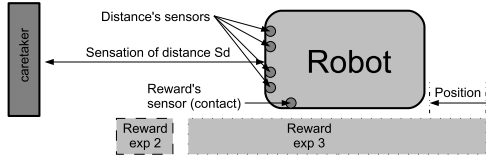
Figure 7: Different positions of the reward.

accelerates (exp 2), and if the reward is lost it tries to come back and stop exploring for a while (exp 2) or forever (exp 3). The robot is able to produce the behaviors shown before (seeking stability and exploration) but can also interrupt these behaviors to take advantage of opportunities and avoid dangers.
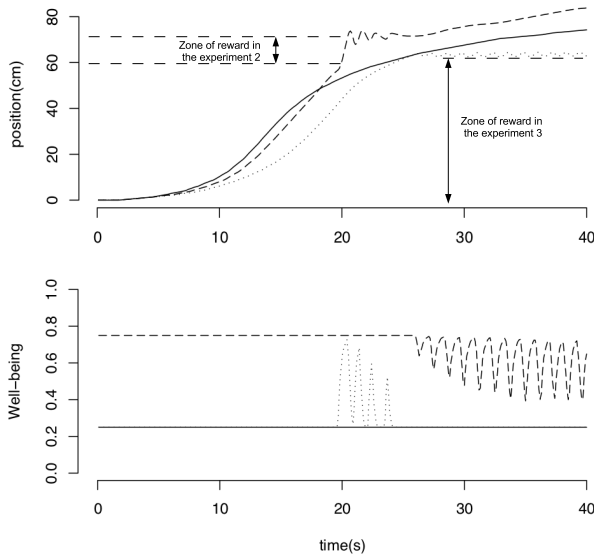

Figure 8: Successive positions of the robot (top) and its corresponding well-being (bottom) for the three experiments—exp 1 without reward in solid line, exp 2 with a local reward in dashed line, and exp 3 with a disappearing reward in dotted line.

Figure 9 depicts the entire architecture allowing the robot to explore when it feels safe, seek stability if it feels uncomfortable, and interrupt these behaviors if it discovers an opportunity or a sudden danger.

5. Low-level Imitation

Imitation can be a means for a robot to learn what is useful (Alissandrakis et al., 2005, Billard, 2000) but its modulation through affect permits to answer other important questions—part of the "big five" according to (Dautenhahn and Nehaniv, 2002)—namely "when" and "what" to imitate. It has been observed in infants that imitative behaviors are closely dependent on comfort (Dunn, 1977, Kugiumutzakis et al., 2005) or affective links with the imitated persons (Hatfield et al., 1994). Interestingly, our current architecture also allows to per-

form low-level imitation as a side effect of the exploratory process depending on affect and pleasure. If the caretaker moves at the front of the robot, the robot receives unexpected sensations and inhibits its exploratory behavior. Moreover, if the robot is not too far from its desired sensation (familiar zone) or its well-being increases, the motivation to continue (Mc) is positive and therefore the robot tries to amplify the new sensations. The result is that the robot moves toward the caretaker when the caretaker moves toward the robot and the robot moves away from the caretaker when the caretaker moves away from the robot. This is not the case if the robot is in an unfamiliar zone or if the well-being decreases. We therefore observe low-level imitation depending on affect and pleasure. This view of low-level imitation differs from other approaches of low-level imitation like (Andry et al., 2003, Demiris and Johnson, 2003) because we do not consider imitation as the reduction of error between what is expected and what is actually sensed but, on the contrary, as the process of amplifying an unexpected or unfamiliar sensation.

Figure 10 presents a typical result of a dynamical interaction with the robot. In this case, we have used the same setup as in the experiment about exploration (see section 3.), but this time the caretaker moved to observe the reaction of the robot. In the top graph we can see the successive positions of the robot in solid line, and the estimated[2] position of the caretaker in dashed line. In the bottom graph see the values (Sd) of the distance sensor sent to the robot. We observe that the robot tries to amplify the relative movement of the caretaker (represented by the arrows in the figure) when its sensation is close to its initial sensation (imitative behavior), but this amplification becomes null or even negative when its sensation is far from its initial sensation (avoidance zone). A robot could therefore imitate another agent to discover new sensations but be able to interrupt this behavior like in the previous section to avoid new dangers or take advantage of new opportunities.

6. Conclusion and Perspectives

We have presented an affect-based PerAc robot architecture implemented and tested in a Koala robot that can appropriately select among a number of behaviors which provide a basis for autonomous and safe learning in an environment that includes a caretaker. Increasing autonomy with respect to the caretaker and exploring the environment are essential activities for autonomous agents to learn new things and to consolidate past experiences and apply them to improve behavior. However, exploration is also

---

[2]Estimation done using the absolute position of the robot and the detected distance of the caretaker to the robot
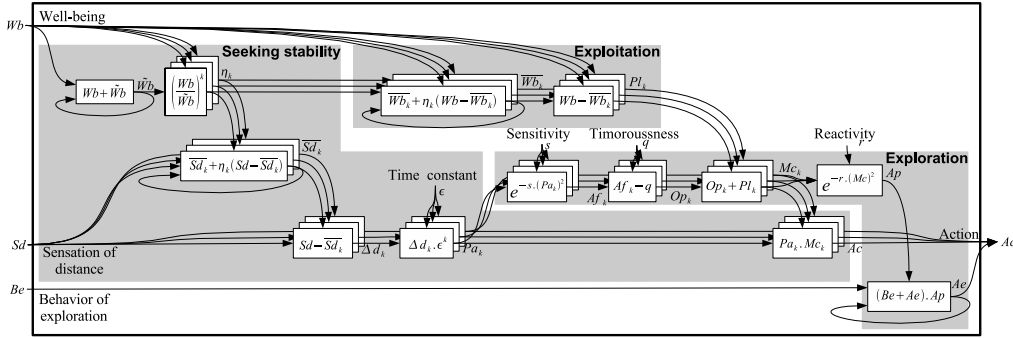
Figure 9: Final architecture allowing to produce seeking stability, opportunism, avoidance and exploration.
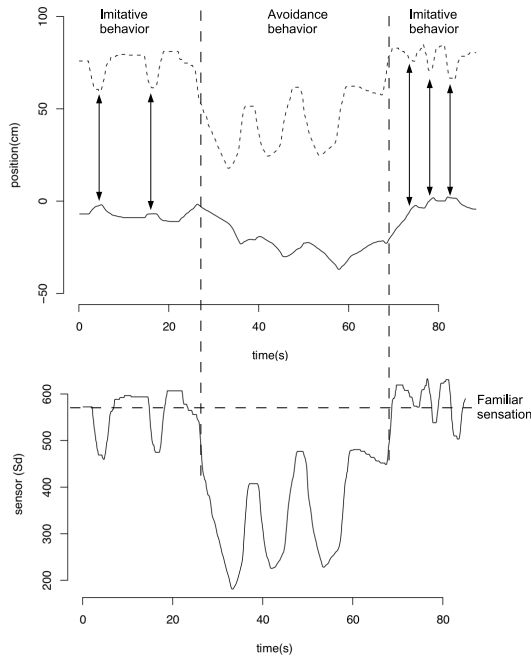

Figure 10: Top: successive positions of the robot (solid line) and of the caretaker (dashed line). Bottom: sensation from the sensors to the robot. When the sensation is close to the familiar sensation (the one the robot starts with), the robot moves like the caretaker (imitative behavior) but when the sensation is far from the familiar sensation, the robot avoids the movements of the caretaker (avoidance behavior).

risky as it exposes the agent to unknown, potentially overwhelming or dangerous situations, and therefore a trade-off must exist between activities such as seeking stability, exploring, imitating another agent (and, in so doing, discovering new sensations) and taking advantage of opportunities offered by new situations and events while at the same time avoiding danger. Our architecture achieves an adapted execution of different behaviors on the grounds of modulatory mechanisms based on notions of "well-being" and "affect." This includes production and modulation of imitative behavior, which is used instead of

the exploratory behavior when a "teacher" proposes new stimulations.

To continue this work, we would like to explore two main avenues. First, we would like to enrich the perceptual space of the robot. At present, we use only one feature (distance to the perceived stimulus) to learn about the "caretaker". Proper treatment of learning about the object of attachment as well as about novel objects would require considering multiple features that the robot would have to analyze in order to recognize the "caretaker" and novel stimuli from different perspectives and in different situations. Second, we would like to further investigate and develop the biological plausibility of our system, in particular its similarity with the brain dopaminergic circuits in their involvement in behavior selection.

Acknowledgments

References

Alissandrakis, A., Nehaniv, C. L., Dautenhahn, K., and Saunders, J. (2005). An approach for programming robots by demonstration: Generalization across different initial configurations of manipulated objects. In *IEEE Computational Intelligence in Robotics and Automation (CIRA2005)*.

Andry, P., Gaussier, P., and Nadel, J. (2003). From sensori-motor development to low-level imitation. In *2nd Intl. Wksp. on Epigenetic Robotics*.

Ashby, W. (1952). *Design for a Brain. The Origin of Adaptive Behaviour*. London: Chapman and Hall, Ltd.

Bateson, P. (2000). What must be know in order to understand imprinting? In Heyes, C. and

Huber, L., (Eds.), *The Evolution of Cognition*, pages 85–102.

Billard, A. (2000). Learning motor skills by imitation: a biologically inspired robotic model. *Cybernetics and Systems*, 32(1-2):155–193.

Blanchard, A. and Cañamero, L. (2006). Modulation of exploratory behavior for adaptation to the context. In Kovacs, T. and J., M., (Eds.), *Biologically Inspired Robotics (Biro-net) in AISB'06: Adaptation in Artificial and Biological Systems.*, volume II, pages 131–139.

Blanchard, A. and Cañamero, L. (2005a). From imprinting to adaptation: Building a history of affective interaction. *Proc. of the 5th Intl. Wksp. on Epigenetic Robotics*, pages 23–30.

Blanchard, A. and Cañamero, L. (2005b). Using visual velocity detection to achieve synchronization in imitation. In Demiris, Y., Dautenhahn, K., and Nehaniv, C., (Eds.), *Third Intl. Symposium on Imitation in Animals and Artifacts*, pages 26–29. AISB'05, SSAISB Press.

Dautenhahn, K. and Nehaniv, C. (2002). The agent-based perspective on imitation. In Dautenhahn, K. and Nehaniv, C., (Eds.), *Imitation in Animals and Artifacts*, chapter 1. MIT press.

Demiris, Y. and Johnson, M. (2003). Distributed, predictive perception of actions: a biologically inspired robotics architecture for imitation and learning. *Connection Science*, 14(4):231–234.

Dunn, J. (1977). Distress and comfort. In Bruner, J., Cole, M., and Lloyd, B., (Eds.), *The Developing Child*, chapter Crying, Comfort and Attachment, pages 67–75. Fontana/Open Books.

Gaussier, P., Moga, S., JP., B., and Quoy, M. (1998). From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence*, 1(7):701–727.

Gaussier, P. and Zrehen, S. (1995). Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16:291–320.

Hatfield, E., Cacioppo, J., and Rapson, R. (1994). *Emotional Contagion*. Cambridge University Press.

Kaplan, F. and Oudeyer, P.-Y. (2004). Maximizing learning progress: an internal reward system for development. In Iida, F., Pfeifer, R., Steels, L., and Kuniyoshi, Y., (Eds.), *Embodied Artificial Intelligence*, LNCS 3139, pages 259–270. Springer-Verlag, London, UK.

Kaplan, F. and Oudeyer, P.-Y. (2005). The progress-drive hypothesis: an interpretation of early imitation. In Dautenhahn, K. and Nehaniv, C., (Eds.), *Models and Mechanisms of Imitation and Social Learning: Behavioural, Social and Communication Dimensions.* Cambridge University Press. to appear.

Kugiumutzakis, G., Kokkinaki, T., Makrodimitraki, M., and Vitalaki, E. (2005). Emotions in early mimesis. In Nadel, J. and Muir, D., (Eds.), *Emotional development*, pages 161–182. Oxford University press.

Likhachev, M. and Arkin, R. (2000). Robotic comfort zones. In *SPIE Sensor Fusion and Decentralized Control in Robotic Systems III*, pages 27–41.

Oudeyer, P.-Y. and Kaplan, F. (2004). Intelligent adaptive curiosity: a source of self-development. In Berthouze, L., Kozima, H., Prince, C. G., Sandini, G., Stojanov, G., Metta, G., and Balkenius, C., (Eds.), *Proc. of the 4th Intl. Wks. on Epigenetic Robotics*, volume 117, pages 127–130. Lund University Cognitive Studies.

Panksepp, J. (1999). *Affective Neuroscience*, chapter 8-SEEKING Systems and Anticipatory States of the Nervous System, pages 144–162. Oxford University Press.

Porr, B. and Worgotter, F. (2003). Isotropic sequence order learning. *Neural Computation*, 15:831–864.

Power, T. (2000). *Play and Exploration in Children and Animals*. Lawrence Erlbaum Associates, Publishers.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9(2):129–154.

Rescorla, R. and Wagner, A. (1972). A theory of pavlovian conditioning: Variations in effectiveness of reinforcement and nonreinforcement. In Black, A. and Prokasy, W., (Eds.), *Classical Conditioning II*, pages 64–99. New York: Appleton-Century-Crofts.

Simon, H. (1967). Motivational and emotional controls of cognition. In *Psychological Review*, volume 74, pages 29–39.

Steels, L. (2004). The autotelic principle. In Fumiya, I., Pfeifer, R., Steels, L., and Kunyoshi, K., (Eds.), *Embodied Artificial Intelligence*, volume 3139 of *Lecture Notes in AI*, pages 231–242. Springer Verlag, Berlin.