Emotion Generation Based on a Mismatch Theory of Emotions for Situated Agents

Clément Raïevsky and François Michaud

Université de Sherbrooke, Canada

Abstract

Emotion plays several important roles in the cognition of human beings and other life forms, and is therefore a legitimate inspiration for providing situated agents with adaptability and autonomy. However, there is no unified theory of emotion and many discoveries are yet to be made in its applicability to situated agents. One function of emotion commonly identified by psychologists is to signal to other cognitive processes that the current situation requires an adaptation. The main purposes of this chapter are to highlight the usefulness of this signaling function of emotion for situated agents and to present an artificial model of anger and fear based on mismatch theories of emotion, which aims at replicating this function. Collective foraging simulations are used to demonstrate the feasibility of the model and to characterize its influence on a decision-making architecture.

Introduction

In spite of significant evidence that emotions play a crucial role in cognitive processes (Frijda, 1986; Hebb, 1949; G. Mandler, 1984; K. R. Scherer, Schorr, & Johnstone, 2001), no consensus currently exists about a unified theory from which an artificial model can be derived. Therefore, implementing emotions in artificial systems can improve our understanding of existing theories and test their usage and effects.

One function of emotion commonly identified by psychologists is to signal to other cognitive processes that the current situation requires an adaptation. This function is very important for situated agents because they operate in a continuous, dynamic and unpredictable world. In this kind of environment, it is crucial to be able to detect events and stimuli, which are relevant to the systems' concerns. This function of emotion is also essential to ensure situated agents autonomy. As human designers cannot anticipate every possible situation a situated agent can come across, complete autonomy is only achievable if situated agents are able to detect situations for which their decision-making mechanism is not adapted. Furthermore, since nobody can fully understand how an artificial agent experiences its reality (Nagel, 1974), the detection of problematic situations should be done without relying on human provided knowledge.

The main purposes of this chapter are to highlight the usefulness of the signaling function of emotion for situated agents and to present an artificial model of anger and fear based on mismatch theories of emotions. The model aims at replicating this function, relying neither on a predictive model of the world nor on specific stimuli analysis.

Background

Psychology is a natural inspiration when designing a model of artificial emotions. However, to clarify its possible contributions to a model, it is useful to classify different theories coexisting. Frijda (1986) proposes such a classification based on three main categories: specific-stimulus theories, intensity theories, and match-mismatch theories.

Classes of psychological theories of emotions

Specific-stimulus theories of emotion state that emotions are triggered by the individual's perception of particular stimuli in the environment. For example, a loud sound, an intense flash, darkness or an unknown object elicit fear. Watson's (1919) theory of emotion belongs to this class of theories. Even if some emotions are actually triggered by this individual's perception, theories belonging to this class do not account for the variety and the complexity of human or apes' emotions. The very same stimulus can elicit very different emotions depending on the internal and external contexts into which this stimulus is perceived. For example, someone losing contact with the ground does not feel the same if he is hiking on a steep slope or if he is diving into water.

According to intensity theories, emotions are associated with different intensities of stimulation or particular variation of this intensity. For example, a gentle contact or a sweet taste are pleasant and associated with positive emotions whereas a strong hit or taste are associated with negative emotions. However, the concept of intensity is not always relevant to stimuli. For example, it is hard to find stimulus intensity related to the fear emotion elicited by an unknown situation. Furthermore, some of these theories include the arousal associated with emotional experience in the intensity, which makes them circular.

According to match-mismatch theories, emotions are elicited by a match or a mismatch between an individual's dispositional entities and events occurring in the environment. Depending on the theories, dispositional entities are called response tendencies, expectations, motives, goals, values, or commitments but they all consider the relevance of the events crucial to the emotion triggering process and especially more important than the event itself. Frijda (1986) includes theories presented by Spinoza (1677), Brown and Farber (1951), Hebb (1949), Mandler (1984), and Lazarus (1966) as belonging to this class of theories. According to Stets (2003), affect control theories (Heise, 1979; Smith-Lovin, Robinson, & Wisecup, 2006) and self-discrepancy (Higgins, 1989) theories also belong to this category. Frijda states that these theories do not focus upon different emotions but rather upon elicitation of emotion as such. He phrases it this way:

> "The notion of 'emotional stimulus' somewhat recedes into the background: Mismatch – interference, interruption, discrepancy – or match – attainment of incentives, correspondence with expectations – is what counts, rather than the precise stimulus that causes that mismatch or match." (Frijda, 1986, p. 267)

According to Brown and Farber (1951), a negative emotion is triggered when a behavior cannot be executed due to either a physical obstacle or the coexistence of conflicting or incompatible response tendencies. Stimuli are emotional if they produce or signal obstruction of a response tendency. In this theory, dispositional entities, which are involved in mismatch, are therefore response tendencies. They echoed the behaviorist origins of this theory. All other theories mentioned by Frijda in the match-mismatch class also belong to cognitive theories of emotion. This approach is characterized by the importance given to cognitive evaluations in emotion generation processes.

Scherer (2001; 1993) states that emotions are elicited from a sequence of evaluations of the stimulus called “Stimulus Evaluation Checks”. This sequence consists of five elements, which are strongly interdependent. Each stimulus is appraised according to:

- its novelty;
- its pleasantness;
- its relevance to current goals;
- the individual’s ability to adapt or react to it;
- its significance regarding social norms and the self concept.

The particular emotion triggered by this sequence depends upon the results of the different evaluation steps. In this theory, the main dispositional entities, which are monitored for match or mismatch with stimuli, are goals.

Similarly, Lazarus (Lazarus & Folkman, 1984; Lazarus, 1999) considers that dispositional entities involved in emotion are goals:

> “*Goal relevance* is fundamental to whether a transaction is viewed by a person as relevant to well-being. In effect, there is no emotion without there being a goal at stake” [his italics] (Lazarus, 1999, p. 92)

Lazarus considers that an emotion is the combination of a cognitive evaluation, an action tendency, and a particular physiological reaction. He states that these different elements are perceived as a single phenomenon. In his theory, emotion differentiation is done through a “secondary appraisal”, which evaluates the individual’s ability to cope with the current “transaction”.

Another cognitive match-mismatch theory is that of Mandler (1984; 1997). Emotions, according to Mandler, are caused by interruptions of behavioral or cognitive activities. He emphasizes the importance of the reaction of the sympathetic nervous system (SNS) to perceived discrepancies in the emotion generation process:

> “The specific relevance of the SNS discrepancy response to emotional phenomena is that SNS arousal is part of the emotional experience and that discrepancies are the major occasions for emotions to occur. I stress that emotional states do not simply occur when the SNS is aroused — the cognitive or evaluative part is a necessary co-condition.” (G. Mandler, 1997, p. 71)

The nature of the interrupted activity determines which particular emotion is elicited. In this theory, dispositional entities, which give stimuli their emotional potential, are action tendencies and goals.

Frijda (1986; 1989) states that emotions are not purely cognitive processes because, according to him, hedonistic evaluations, which play a crucial role in emotions, are not cognitive and changes in one’s action readiness associated with some emotional states take place unintentionally. While Frijda does not consider that emotions are only cognitive, his theory clearly belongs to cognitive theories. In effect, the emotion process he sets out (Frijda, 1986) includes various appraisals of the stimulus event and of the stimulus situation. In particular, the relevance of the stimulus event for one or more of the individual’s concerns is evaluated and the possibilities of coping are assessed.

Conclusions on classes of theories regarding their usefulness for situated agents

Specific-stimulus theories account for the part of emotion that associates particular configurations of the environment with predefined reactions. This subset of emotion-related processes is useful to improve individual's reaction to situations requiring a quick behavior adaptation. Specific-stimulus theories are therefore interesting to enhance specific reactions of artificial systems to a set of predefined situations. However, artificial models of emotions based upon this kind of theories are fundamentally limited by the fact that a human designer cannot fully grasp the reality of a situated agent (Nagel, 1974). Practically, the set of situations a designer can anticipate is limited whereas a situated agent can face infinity of situations. Furthermore, associating specific reactions with particular configurations of the environment denotes a specialization to the mission of the system. Models inspired by specific-stimulus theories are therefore specific to a set of situations and to the mission of the agent for which they are designed. Such models are consequently limited in the adaptability they can provide to situated agents.[1]

Intensity theories of emotion are even more specialized to a particular set of situations because they rely on the existence of a measurable intensity in some feature of the environment. Furthermore, these theories do not account for many aspects of emotion. The limitations of artificial models of emotion based on specific-stimulus theories apply also on models inspired by intensity theories.

Like Frijda, we believe that match-mismatch theories are a more fruitful avenue to explain the richness of emotions than the specific-stimulus one. Frijda's argument to support this viewpoint is that:

> "Emotions are rarely, if ever, elicited by an isolated stimulus. Rather, the emotional effectiveness of sensory stimuli depends upon the spatial, temporal, and meaning context in which they occur, the adaptation level upon which they impinge, and the expectations with which they clash or correspond." (Frijda, 1986, p. 267)

Match-mismatch theories of emotion are strongly related to the signal function of emotion because they state that emotion occurs when a match or a mismatch is detected. Therefore, the elicitation of an emotion intrinsically signals the occurrence of a situation requiring one's attention and adaptation of one's behavior. We are in agreement with Sloman on the importance of the signaling function in emotion processes when he states:

> "The main function of the mechanisms referred to above [(emotions)] is to prevent 'normal' processing from continuing in circumstances where some state requiring (or prima-facie requiring) a change of 'direction' occurs." (Sloman, 2008)

As we stated before, this function of emotion is crucial to situated agents because it allows them to detect and respond to situations to which their decision-making mechanism is not adapted. This ability is essential to trigger adaptations in the decision-making mechanism when required. An important characteristic of this function is that it is not specific to a particular set of situations or to the current objectives of the individual. Therefore, artificial models of emotion inspired by match-mismatch theories must not be specialized to particular configurations of the environment or to the mission assigned to the agent.

---

[1] One can argue that the versatility of the system can be improved by adding analyses of increasing complexity to the stimulus evaluation. However, the more complex the analyses are, the more resources consuming they are. Once a particular level of complexity is reached, they cannot be conducted continuously in a complex environment.

Emotions in Artificial Intelligence (AI)

Emotions have inspired researchers since the very beginning of AI. For instance, Simon (1967) stated that artificial systems must have an interruption mechanism comparable to the role of emotion. According to him, artificial models of human intelligence should include two main sub-systems: one in charge of carrying out the system's goals and one monitoring the environment and the system itself, which is able to interrupt the first sub-system when the situation requires an adaptation. Simon asserts that this second sub-system corresponds to emotion processes.

Theories of emotion proposed by Ortony, Clore and Collins (1988) and by Oatley and Johnson-Laird (1987), while being deeply rooted in psychology, are explicitly formulated to make it possible to implement in an artificial system. Unlike Simon who emphasizes the function of emotion and considers the emotion generation process, Ortony, Clore and Collins' theory focuses upon the differentiation of emotions. According to them, emotions are differentiated from an appraisal process, which evaluates events occurring in the environment, actions of other agents, and perceivable objects of the environment. This appraisal process is composed of two levels: first, events are appraised as pleasant or not, action of agents are endorsed or rejected, and objects are evaluated as attractive or repulsive. Once an event has been characterized as relevant, the agent to which it is relevant is assessed. If the event is relevant for another agent, its consequences are appraised as desirable or not. If the event is relevant for the agent itself, its consequences are evaluated as relevant or not to the concerns of the agent. Consequences of events relevant to the agent's concerns are further differentiated between confirmed and disconfirmed consequences.

The basic assumption of Oatley and Johnson-Laird is that the human cognitive system is composed of modular and asynchronous processes organized as a hierarchy. Each process has a goal, which it attempts to reach if certain preconditions are met and this, without interruption. According to Oatley and Johnson-Laird, the main role of emotions is to coordinate the different processes making up the system. Emotions do so by setting the whole system in one of a small number of "emotion modes" through a non-propositional communication channel. These modes allow the system to give specific priorities to processes in the hierarchy and to organize transitions between plans.

Both theories we have just briefly described are explicitly aimed at reducing the gap between psychological concepts related to emotion and possible implementation in an artificial system. As such they both have had a great influence on models of emotion designed for artificial systems.

Artificial emotions for situated agents

Situated agents designers have taken into account psychological theories of emotion in two main ways (Scheutz, 2004): on one hand, "effect models" of emotions only implement overt and observable aspects of emotional behavior, regardless of the underlying processes. On the other hand, "process models" of emotions try to replicate internal processes and functions related to emotion. These later models are more deeply rooted in psychological theories than effect models which aim at making the observer believe that the system is experiencing emotions.

Various emotional mechanisms related to process models of emotions have been implemented in situated agents to enhance the quality of interaction between humans and synthetic agents (Velásquez, 1998), to increase synthetic agents learning abilities (Gadanho, 2002), or to improve coordination between situated agents (Murphy, Lisetti, Tardif, Irish, & Gage, 2002; Parker, 1998). These last two approaches are the most relevant to our work.

Murphy et al. (2002) set out a two-layered control architecture with one sensory-motor level and one schematic level. The sensory-motor level is responsible for controlling the agent through modular behaviors and for detecting specific features or events in the environment. The schematic level controls behaviors and organizes them into more abstract abilities. This later level is implemented by two finite state machines: a behavioral states generator and an emotional states generator. Emotions are represented in this later finite state machine by four states (Happiness, Confidence, Concern and Frustration) and are derived from tasks progress evaluation. This evaluation takes into account features extracted by the sensory-motor layer and messages received from other agents. This approach therefore echoes principles of specific-stimulus theories of emotion. The emotional states generator directly modifies behavior parameters and triggers specific behaviors. This artificial model of emotions is then specific to the environment and the mission of the agents for which it has been designed.

Parker's ALLIANCE architecture (Parker, 1998) is a behavior-based architecture, using motivations to activate different groups of behavior producing modules. In each group, behaviors are arbitrated using subsumption (Brooks, 1991), a priority-based arbitration scheme. Emotions are not explicitly identified in the architecture. However, Impatience and Acquiescence are two motivations serving as temporal measures of task progression. Impatience enables an agent to handle situations when other agents fail in performing a given task, while Acquiescence is useful when the agent itself fails to properly perform its task. Motivation intensities are generated from an analysis of the current agent perception and from dedicated social messages. As for Murphy et al (2002), this approach corresponds to the specific-stimulus theories of emotion. Moreover, each motivation is directly associated with a set of behaviors, which are activated when the motivation intensity is greater than a threshold. These motivations are therefore tightly coupled with the agents' mission objectives.

To summarize, process models of emotions for situated agents are generating emotions on the basis of specific-stimulus theories, confining agents to specific and therefore limited environmental conditions. In addition, these models associate emotional states with particular behavioral responses related to the agent's mission objectives, making them even more specific. Consequently, previous process models of emotions designed for situated agents are limited either because emotion generation is too specifically related to the environmental conditions or because emotional responses are too tightly associated with the agent’s mission objectives; if the environment or the mission changes, the models of emotions become invalid, and this should not be the case in order to capture the true versatile nature of emotion. Emotions should rather be derived from a generic model to capture the fact that different situations can lead to the same emotion, and that the same situation can lead to different emotions.

The infrequent use of match-mismatch theories in situated agents could be due to the difficulty to evaluate the relevance of events with regard to the agent’s concerns, especially in a continuous, dynamic and unpredictable environment. Solving this issue requires at least two abilities: an agent must be able, first, to identify significant events based on its perception and, second, to evaluate the “distance” between desired and actual states of the world. These abilities are strongly linked with the way a situated agent experiments its reality.

Therefore, it will always be difficult to implement such abilities in artificial systems because the knowledge of how an agent experiments its reality is not accessible to human designers. To address this fundamental problem, these abilities could be learned; however this is a complete research avenue, which at the time being, we have chosen not to address. Instead, we have decided to detect mismatches from the monitoring of temporal models of internal resources use. The purpose of our research is to validate the feasibility of an artificial model of anger and fear inspired by match-mismatch theories of emotion. This inspiration entails that, contrary to previous work, our model must not be specific to a set of particular situations or to the mission assigned to the situated agents for which it is designed.

Artificial model of anger and fear

Our model aims at replicating functions of emotion identified in match-mismatch theories, such as detecting and highlighting situations representing a mismatch between the agent's concerns and the experienced condition in the world. This process allows an agent to trigger an emotional adaptive reaction when regular decision-making processes are no longer valid. The emotion generation associated with this adaptive reaction is made independently from environmental conditions by analyzing the agent's intentions in relation to its current intended actions. The specific reactions triggered by emotional experiences are realized by dedicated processes, making our model of emotions independent of the agent's mission objectives.

Figure 1 illustrates how our model works in relation to elements associated with the decision-making processes of the agent. More specifically, the agent must have cognitive processes (Motivations) responsible for generating intentions and determining their desirabilities. Actions of the agent have to be realized by concurrent processes (Behaviors) which are activated by an Action Selection mechanism according to the agent's intentions. These intentions must be explicitly represented.

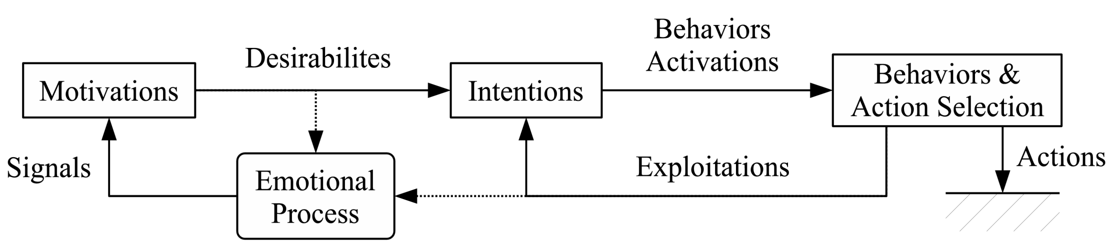

Figure 1 Representation of the cognitive processes required with our model of emotions

Using this model, the Emotional Process looks at how Intentions are carried out in order to monitor their compliance with the agent's Motivations. When a mismatch is detected by the appearance of a discrepancy between the agent's Intentions and the way they are satisfied by the Exploitation of behaviors, the Emotional Process tries to identify the source of the discrepancy by looking at the agent's current Intentions. The occurrence of this mismatch and its cause are then signaled to the agent's Motivations, which can change the agent's Intentions accordingly. Like match-mismatch theories of emotion, this artificial model is not focused on which particular emotion is elicited in predetermined situation but rather on getting the emotion process under way.

Monitoring of discrepancies between the agent's Intentions and behavior Exploitation is done through temporal models. These models depend on the type of Intentions: Goal-Oriented Intentions (GOI) are related to behaviors that must be exploited to make the agent accomplish tasks aimed at fulfilling the agent's goals. In favorable situations, behaviors associated with these Intentions are exploited when activated. This is reflected in the satisfaction condition of GOI, which is to be exploited when desirable. Conversely, Security-Oriented Intentions (SOI) keep the agent away from problematic situations and behaviors associated with them are not exploited in favorable situation. Satisfaction condition of SOI is therefore not to be exploited when desirable.

Models equations

The accumulated time $a_I(t)$ during which intention I does not conform to its satisfaction condition at time t is expressed by (1) and (2):

$$a_I(t) = \int_{-W_I}^{0} b_I(t)dt \quad (1)$$

$$b_I(t) = \begin{cases} 0 & \text{if intention } I \text{ conforms to its satisfaction condition at time } t. \\ 1 & \text{else} \end{cases} \quad (2)$$

where W_I is the length of the sliding time window over which intention I is monitored. A mismatch is detected when $a_I(t)$ becomes greater than a time threshold T_I.

In order to replicate the signaling function of emotion, once a mismatch is detected, its cause has to be determined and highlighted. The cause of the mismatch is found out through an analysis of the current agent's intentions. To understand this analysis, it is important to distinguish the intention **triggering** the mismatch (by not conforming to its satisfaction condition) from the intention **causing** this mismatch (by preventing the triggering intention to conform to its satisfaction condition). As depicted in Figure 2, a mismatch related to a Goal-Oriented Intention is detected if it is desirable but not exploited during a certain period of time. Therefore, the intention identified as being the cause is the one being exploited the most during the recent past, because it hinders the exploitation of the triggering intention.

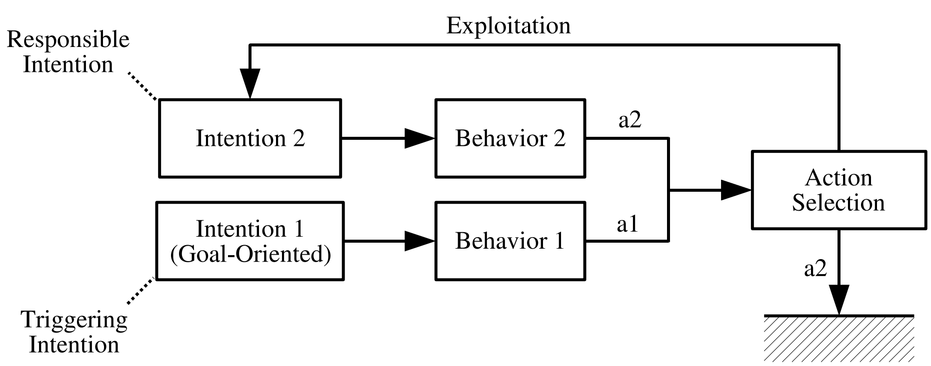

Figure 2 Mismatch detection involving a Goal-Oriented intention

Conversely, a Security-Oriented Intention triggers a mismatch when it has been exploited for a certain period of time; in this case, the cause is the Security-Oriented Intention itself. Once identified, the cause is highlighted by signaling the concerned intention to the agent's motivations. Motivations can then use this information to change the intentions of the agent in order to adapt its behavior to the situation.

Emotions differentiation

Once a mismatch has been detected, it is necessary to determine which particular emotion will be elicited. Works in psychology and ethology (Lazarus, 1984, 1999; Weiss 1971) have shown that in stressful situations, the elicited emotion depends on one's belief in one's ability to handle this situation.

Discussing fearful situations, Lazarus (1984) writes:

> "As efficacy expectancies increase and the person judges his or her resources more adequate for satisfying task demands, the relationship is appraised as holding the potential for more control and therefore as less threatening. As a consequence, fear level decreases and coping behaviors are instituted." (p. 70)

Frijda (1986) expresses a similar viewpoint more concisely: "Fear is the emotion of uncertainty and lack of control." (p. 429). Conversely, in situation eliciting anger, an individual can still modify its relation with its environment; control has not been lost. When discussing such situations, Frijda states that "behind the obstacle, the blocked goal still exists, still is available; and the nature of the obstacle is such that, in principle it can be controlled and modified." (p. 429)

In our model, emotions are differentiated according to the nature of the triggering intention. This choice is in accordance with Frijda and Lazarus positions about stressful situations because the nature of the triggering intention reflects the agent's belief in its ability to control the situation. Security-Oriented intentions exploitation model states that these intentions should not be exploited. So if a Security-Oriented intention does not conform to its temporal model, this is because it is exploited, i.e., the agent 'believes' it is carrying out the actions associated with the triggering intention. From the agent's viewpoint, its actions have no effect on the environment.

Therefore, when a Security-Oriented intention triggers a mismatch, it is the Fear emotion variable that is activated. Conversely, when a Goal-Oriented intention triggers a mismatch, the agent considers it is not carrying out the behaviors associated with this intention. It therefore determines that it still has the ability to modify the situation. In this case, it is the Anger emotion variable that is activated.

Emotion variables behavior

Each detected mismatch contributes to the activation of an emotion variable. All these contributions are summed to determine the emotion variable's intensity $E(t)$. This intensity is expressed by Eq. (3):

$$E(t) = \sum_{\forall I} c_I(t) \qquad (3)$$

where $c_I(t)$ is the contribution of the mismatch triggered by intention I to the intensity of the emotion variable. As long as the intention I is involved in a mismatch, $c_I(t)$ increases linearly with time.

As soon as the mismatch involving I disappears, $c_I(t)$ decreases according to a function described by Picard (1995), which corresponds to the decrease of the influence of a stimulus on an emotion. $c_I(t)$ is written in the form:

$$c_I(t) = \begin{cases} c_I(t_0) + A_I \cdot (t - t_0) & \text{if } a_I(t) > T_I \\ c_I(t_1) \cdot e^{-D_I \cdot (t - t_1)} & \text{else} \end{cases} \tag{4}$$

where A_I is a constant increasing factor associated with intention *I*, D_I a constant decreasing factor associated with intention *I*, t_0 the instant when $a_I(t)$ has become greater than T_I and t_1 the instant when $a_I(t)$ has become lesser than T_I. Intensity of emotion variables are kept between 0 and 1.

Motivations can use the current emotional state provided by the emotional process to modulate the agent behavior parameters. For example, an angry agent can spend more energy in its behavior, as suggested by Piaget (1989).

Interestingly, using an intention-exploitation model makes our emotional process independent of the way intentions are carried out by the agent. Furthermore, these models are independent of the situation experienced by the agent because sensory data are not taken into account. Emotional experiences are therefore generated independently of both particular environmental conditions and specific behaviors. The model of emotions itself is uncoupled from the agent's mission objectives because reactions specific to the highlighted problematic situations are carried out by motivations. This independence between our model and, first, particular situations and, second, the agent's mission ensures that our emotional process is relevant to other applications and respond to the constraints entailed by the inspiration from match-mismatch theories of emotion.

Experiments

Multi-agent foraging (Goldberg & Matarić, 2002) is a widely used task to investigate group behavior in constrained and dynamic settings, with clear metrics to evaluate performance (i.e., physical interferences, traveled distance, time to complete). This explains our decision of using such a task to illustrate the working of our emotional process. Fig. 2 illustrates the simulated environment (implemented in Stage (Vaughan, Gerkey, & Howard, 2003)) used for the foraging experiments. The simulated agents are Pioneer 2 DXs in a pen of 6 × 10 meters. Six agents have to collect 12 pucks and take them onebyone to the home region. Each agent is given two simulated sensors: one laser range finder with an 8 meter-range and a 180° field of view, and one fiducial finder, returning the identifier and relative position of objects with a fiducial tag, in a range of 5 meters and a 180° field of view. Each agent has a unique fiducial identifier, which allows it to perceive others' relative positions. Home flags and pucks also have fiducial ids. Agents are equipped with a gripper allowing them to collect one puck at a time. Note that such simulated conditions are realistic since they are based on the existence of real sensors, such as laser range finder, color cameras for pucks and home detection, and omnidirectional ultrasonic positioning devices (Rivard, Bisson, Michaud, & Létourneau, 2008) for localizing and identifying agents. Agents are considered to be homogeneous as they all have the same physical and decisional capabilities. They can communicate with each other using broadcast mode (through network link).

The group coordination strategy is based on a dominance hierarchy: when an agent perceives a higher-ranked agent in a range of 1.5 meter in front of it, it stops. This range is called the social range of the agents and has been dynamically changed according to the agents' emotional states. This distributed strategy aims at avoiding physical interference while minimizing traveled distance. The choice of a hierarchy to coordinate a group comes from the social function of emotion and especially from the work of Plutchik (1980) who states that emotions of fear and anger play an important role in the dominance hierarchy, which is established in ape groups and in some part of the human society. He also states that these dominance hierarchies are useful for social groups since they lower the frequency of conflicts for shared resources such as food or mate.

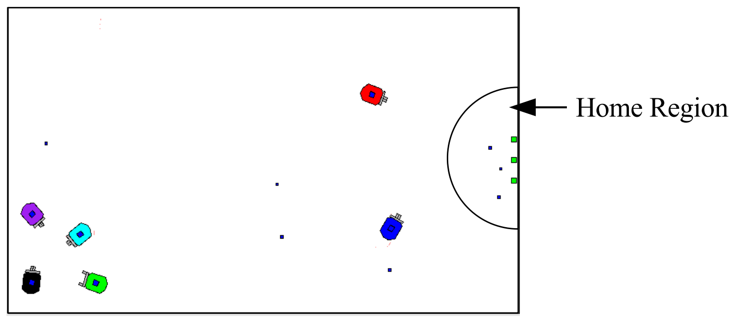

Figure 3 Simulated experimental settings

Implementation

The framework used to implement our emotional process is based on a distributed behavior-based system named MBA (Motivated Behavioral Architecture) (Michaud et al., 2007). Figure 3 illustrates the MBA architecture with the specific modules implemented for the collective foraging task. Each agent in the group uses this architecture for decision-making.

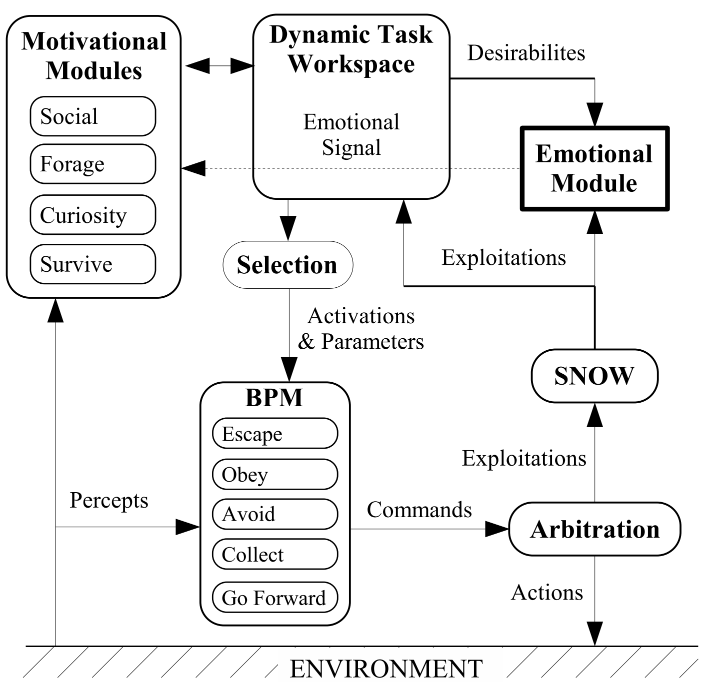

Figure 4 Implementation of MBA for foraging agents

Behavior-producing modules (BPM, or behaviors) constitute basic components from which the agent can operate in the world. BPM are configured to issue commands based on the agent's perception and intentions. For the collective foraging task, the five BPM are arbitrated using subsumption in the following order of priority:

- Escape, which makes the agent turn on itself to find a safe passage to leave the current location.
- Obey, which makes the agent execute a particular action such as stopping or turning left, according to a parameter associated with the agent's intentions.
- Avoid, which makes the agent move safely in the environment by avoiding obstacles using laser range finder readings. Only obstacles within a 0.9 meter radius of the agent are taken into account.
- Collect, which tracks pucks, collects them one at a time and takes them back to the home region.
- Go Forward, which gives the agent a constant linear velocity.

These BPM issue commands only if they are activated. Their activation and parameters are derived by the *Selection* module based on the agent's intentions. These intentions are generated by the *Motivational Modules* (MM) and are stored in the *Dynamic Task Workspace* (DTW). They are organized in a tree-like structure according to their interdependencies, from high-level/abstract intentions to primitive/behavior-related intentions. For example, a high level intention could be Stay Safe and its children in the hierarchy could the Avoid and Escape intentions which are associated with behaviors.

MMs are asynchronous, independent modules that can add intentions, modify or monitor their parameters, and give recommendations about them. These recommendations correspond to the desirabilities of intentions according to MMs and can take three different values: positive, negative and undetermined. The *Selection* module applies a policy to these recommendations to determine the behaviors' activations; a behavior is activated if its corresponding intention has at least one positive and no negative recommendation. The associations between lowest, behavior-related, intentions and behaviors are implemented in the *System Know- How* (SNOW) module. This module is also in charge of updating information about exploited intentions, that is to say intentions carried out by current agent's actions. For the multi-agent foraging task, the Motivational Modules and the intentions related to them are:

- **Curiosity**, which makes the agent explore its environment by recommending the *Explore* intention associated with Go Forward.
- **Forage**, which handles the foraging task by recommending the *Forage* intention (associated with the Collect behavior) and inhibiting the *Stay Safe* intention when a puck is about to be collected by the agent.
- **Survive**, which ensures the security of the agent by adding and recommending the high-level *Stay Safe* intention, and specifying it by adding *Avoid* intention or *Escape* intention as its child. These two intentions are directly associated with behaviors with the same names. We describe later how and when the **Survive** MM chooses between *Avoid* and *Escape*.
- **Social**, which carries out the group coordination strategy through recommendations of the *Obey* intention. This intention is associated with the Obey behavior.

The Emotion Module (EM) implements the model of artificial anger and fear described above. The behavior modifications triggered by the emotions for the foraging task are of two kinds. First, the current emotional state is used to modulate parameters of the agent's behavior. Second, signaled mismatches are used to adapt its behavioral strategies. The **Survive** MM uses the current emotional state to modify the rotation speed of the Avoid behavior and the mismatches detection signals to switch between two avoidance strategies. The **Social** MM uses the current emotional state to change the social range of the agents and uses the mismatches detection signals to adapt the coordination strategy by updating the group's dominance hierarchy.

The Avoid speed rotation is changed from 20 mm/s to 40 mm/s according to the intensity of the Anger variable if it is the dominant emotion, and is reset to 20 mm/s when the dominant emotion is Fear. This influence of the emotional state on the agent's behavior has been chosen according to the position of Piaget (1989) stating that emotions regulate the *intensity* of behavior. The rotation speed of the Avoid behavior can be seen as its own intensity since, in our implementation, this behavior only gives rotation commands to the motor system of the agent. Therefore, an angry agent, having difficulty achieving its Goal-Oriented intentions, will be more energetic with its avoidance behavior to give it a chance to succeed, while a fearful agent will move more carefully because it considers its Avoid behavior has become ineffective in the present situation.

The avoidance strategy is switched from simple obstacle avoidance to escape when the *Stay Safe* intention is signaled as the cause of a mismatch. This is done by the **Survive** MM which replaces the *Avoid* intention with the *Escape* intention as the child of *Stay Safe* in the intentions hierarchy. This change causes the Avoid behavior to be inhibited and the Escape one to be activated.

The social range of an agent varies according to the current emotional state of the superior to which it obeys. An agent stays farther from angry superiors than from superiors in a neutral emotional state and can come closer to fearful superiors. The social range of an agent is expressed by Eq. (5):

$$S(t) = \begin{cases} S_0 \cdot (1 + E_{Anger}^{Sup}(t) \cdot K_s) & \text{if Anger is the superior's dominant emotion} \\ S_0 \cdot (1 - E_{Fear}^{Sup}(t) \cdot K_s) & \text{else} \end{cases} \quad (5)$$

where S_0 is the default social range (1.5 meters in the experiments), E_e^{Sup} is the intensity of emotion e experienced by the superior agent, and K_s the social range variation factor. This factor is set to 0.2 in the experiments, allowing $S(t)$ to vary between 1.2 and 1.8 meters. Values of these parameters have been fixed according to the distances at which agents start to interact physically, beginning to use their avoidance behavior due to other agents nearby.

The coordination strategy is adapted when the *Obey* intention of an agent is signaled as the cause of a mismatch. This typically occurs when an inferior agent obeys to a superior, which is experiencing some kind of failure. When the **Social** MM receives a signal from the Emotion Module meaning that the *Obey* intention it has added is the cause of a mismatch, it triggers an update of the group hierarchy. This update involves only the two agents concerned by the *Obey* intention: the inferior agent and the superior to which it obeys. The update process begins with the inferior sending a message to the superior, challenging the position of the superior in the hierarchy. The answer of the superior depends on its current emotional state: if its dominant emotion is Fear (meaning it is in an unwanted situation involving its security), the superior will resign and leave the inferior's message unanswered.

The superior will thus drop in the hierarchy from the inferior's viewpoint and vice versa. As a consequence, if there are communication problems between agents experiencing situations requiring hierarchy adaptations, they will end up considering all other agents as inferiors. Hence, if communication between the group members fails, their collective behavior will tend gradually to a non-coordinated behavior. Conversely, if the superior is experiencing Anger, it will reply to the inferior and the hierarchy will remain the same. In this case, the update process is reinitiated every 7 seconds as long as the *Obey* intention of the inferior agent is signaled as the cause of a mismatch. When a superior agent accepts to drop in the hierarchy, it does not reply to the inferior's request.

To better understand the hierarchy adaptation process, let us consider the situation depicted in Figure 3. In this situation, the black agent in the bottom left corner is unintentionally surrounded by inferior agents. These agents are motionless because they are obeying to the superior one and are thus blocking it. Figure 5 presents the decision-making mechanism status of one of the inferior agents in this situation, and Figure 6 shows one of the inferior agents' evolution of internal variables during the hierarchy adaptation process. As the problematic situation goes on, from time 0 to time 75, $a_{Forage}(t)$ increases because the *Forage* intention is not exploited. When $a_{Forage}(t)$ reaches the *Forage* time threshold (T_{Forage} = 60 sec.), a mismatch is detected. This detection triggers an analysis of the current intentions, with conclusion that the *Obey* intention was exploited the most in the recent past and is therefore the cause of the mismatch. This information is communicated to the motivations. In response to this signal, the **Social** MM, which is responsible for the addition of the *Obey* intention, triggers the hierarchy adaptation process. As the superior agent is blocked, it is (around time 175) experiencing fear and therefore will not answer to the inferior's request to move down the hierarchy, causing a switch of position between itself and the inferior agent. Once this switch is completed, the former inferior agent has no more reason to stop and can resume foraging.

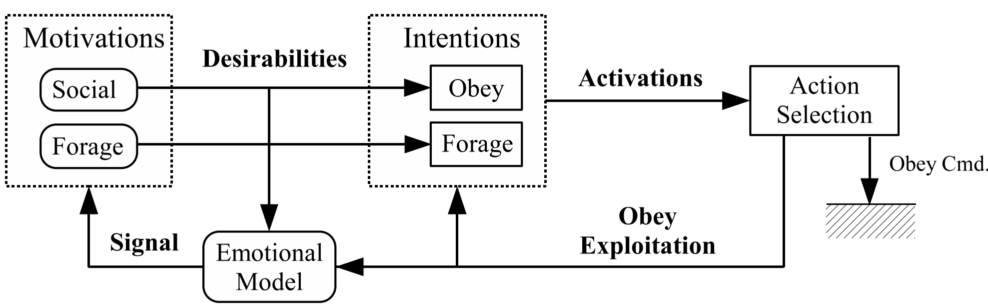

Figure 5: Intentions and behaviors of an agent obeying to a superior experiencing some kind of failure

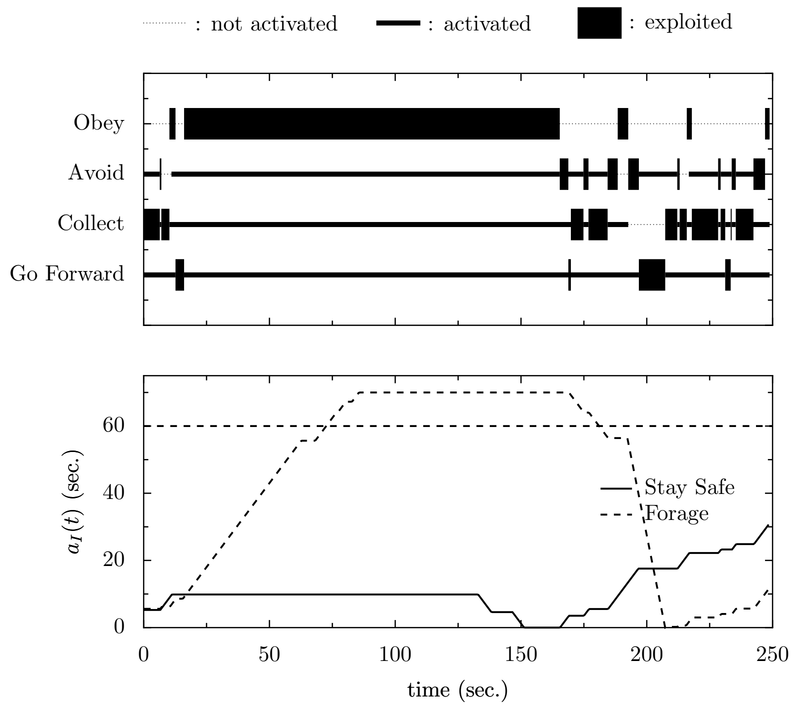

Figure 6: Exploitations and activations of behaviors (top) and accumulated time during which *Stay Safe* and *Forage* intentions do not conform to their exploitation models (bottom) of an agent obeying a superior agent, which is experiencing a failure

Experimental results

To characterize the influence of our emotional process while neutralizing the influence of other architectural components, we conducted a series of trials in each of the following conditions:

- Neither emotional process nor coordination (NE-NC). The **Social** MM or the EM are not used in the decision-making architecture. This condition is used as a reference to evaluate the influence of the emotional process and of the coordination strategy.
- Modulation of the behavior parameters according to the dominant emotion intensity and no group coordination (E_I-NC). This condition is used to characterize the influence of the emotional modulation of behavior parameters.
- Adaptation of strategies triggered by mismatch detection and no group coordination (E_M-NC). This condition measures the influence of switching between avoidance strategies (i.e., between *Avoid* and *Escape* intentions).
- Combination of the two previous conditions and no group coordination (E_{MI}-NC).
- Group coordination without emotional influence (NE-C). The **Social** MM is used but not the EM. The dominance hierarchy is static and predetermined (Goldberg & Matarić 1997). This condition demonstrates the performance of the coordination strategy and is used as a reference to evaluate the influence of the emotional process on the coordination strategy.
- Group coordination and modulation of the behavior parameters according to the dominant emotion intensity (E_I-C).
- Group coordination and adaptation of strategies triggered by mismatch detection (E_M-C).

- Complete emotional mechanism and group coordination (E_{MI}-C). This involves all the emotional processes of our approach.

Each of these conditions has been tested on the same series of forty randomly-generated configurations of agents and pucks in the simulated pen to attenuate the influence of the initial positions. In these trials, the parameters of Eq. (4) are the following: $A_I = 0.7$ and $D_I = 0.025$. These values were chosen to make emotional experiences duration coherent with intention models parameters. *Forage* is a Goal-Oriented intention with $W_I = 70$ seconds and $T_I = 60$ seconds. This intention is considered as satisfied by the agent behavior when the Collect or the Move Forward behaviors are exploited. *Stay Safe* is a Safety-Oriented intention with $W_I = 140$ seconds and $T_I = 120$ seconds.

Table 1: Experimental results

	No Group Coordination				Group Coordination			
	NE-NC	E_I-NC	E_M-NC	E_{MI}-NC	NE-C	E_I-C	E_M-C	E_{MI}-C
Completion Time (sec.)	438	427	392	449	345	406	562	509
Traveled Distance (m.)	391	381	357	396	295	326	398	377
Physical Interference Ratio	24%	25%	24%	24%	13%	13%	16%	15%
Success Rate	90%	92.5%	97.5%	90%	57.5%	52.5%	92.5%	100%

Table 1 summarizes the observed results in terms of the following metrics. Completion Time is the time spent to take all pucks to the home region. Traveled Distance is the total distance traveled by the agents during one trial. Physical Interference Ratio is the amount of time spent at a distance of 0.7 m from other agents. The interference distance is set at the distance at which the linear speed of the agents equals zero when the Avoid behavior is exploited. Physical Interference Ratio echoes the risk of collision between agents and has proved to be a good measure for a coordination strategy in a foraging task (Goldberg & Matarić, 1997). Success Rate is the ratio of failed trials (determined when pucks remain to be collected after 30 minutes) over the total number of trials. This metric captures the ability of the group to recover from situations causing it to fail and we have not anticipated, reflecting the adaptability of the group. All these metrics, except Success Rate, take only successful trials into account not to be offset by the failure time which was set arbitrarily.

The following observations can be made from these experimental results. For conditions without group coordination, modulating behavior parameters according to emotion intensities (E_I-NC) or triggering switching of avoidance strategies from mismatch detection (E_M-NC) resulted both in a higher success rate compared to the reference condition (NE-NC). This suggests that emotional influences improved the agents' adaptability at the individual level. Completion Time, Traveled Distance and Success Rate of the E_{MI}-NC condition indicate however that the influences of mismatches detection and emotions intensities interfere with each other. The dynamic of this interference remains to be fully understood. It seems that the modification of the Avoid rotation speed prevents the strategy switch between Avoid and Escape behaviors to be beneficial to the agents. In conditions without coordination, agents always operate as individuals, and consequently similar performances are observed in terms of Physical Interference Ratio for these experimental conditions.

As expected, the introduction of group coordination strategies has clearly reduced the physical interferences between agents, making them safer. However, modifying behaviors parameters according to emotions intensities to the coordination strategy (E_I-C) has actually decreased performances of the group for the other metrics, compared to NE-C conditions. The increase in Completion Time of the E_I-C condition can be explained by the fact that, in this condition, inferior agents are influenced by farther superior agents than in NE-C condition, making them spend more time motionless. Near the end of the mission, when all pucks have been picked up but not yet brought back to the home region, these delays cause the agents not carrying pucks to travel more distance than in the NE-C condition. That is why Traveled Distance of the E_I-C condition is slightly greater than the one of the NE-C condition.

Combining coordination with strategy modifications based on mismatches detection (E_M-C) drastically improves the Success Rate of the group while keeping the Physical Interference Ratio low. However, this improvement has a cost: this condition has the longest Completion Time. This is implicitly caused by the use of temporal intention models. For example, when an agent carrying a puck is experiencing a situation from which it can resume only through the detection of a mismatch, this causes a delay in the completion of the mission. Furthermore, with the chosen coordination strategy, this agent causes inferior agents nearby to stop, hindering them in the execution of their task and thus further delaying the success of the mission. For instance, in the situation illustrated by Figure 3, all inferior agents involved remain still until the hierarchy update. The Traveled Distance and Physical Interference Ratio of the E_M-C condition are worse than the ones of the NE-C and E_I-C conditions. However, as failed trials are not taken into consideration in these conditions (Success Rate being around 55%), negative effects of problematic situations are not reflected in their metrics. Their low success rates have two main causes: first, when an agent fails by getting stuck on an obstacle or by hitting another agent, all its inferiors passing nearby stop moving, potentially leading to a failure of the group. Second, the situation illustrated by Figure 3 creates a deadlock and leads to a group failure.

The same observations can be made about the comparison of the E_{MI}-C condition with NE-C and E_I-C conditions. In addition, E_{MI}-C condition achieved the best Success Rate and a better Traveled Distance than E_M-C. This shows that taking emotion intensities into account is meaningful for the group only if mismatch detection also influences the group behavior.

Overall, the coordination strategy effectively improves agents' safety by lowering their physical interactions, but needs to be adapted in order not to compromise the success of the group's mission by creating deadlock situations. Taking only emotion intensities into account do not results in significant improvements at the individual level, and decreases the performances of the group. Modifying avoidance and coordination strategies according to mismatch detection improves the agents' ability to recover from unforeseen problematic situations, which cause their decision-making mechanism to fail. Mismatch detection also allows agents operating without coordination to improve their Completion Time and Traveled Distance. Coordinated agents using this mechanism (E_M-C) keep their Traveled Distance and Physical Interference Ratio low, but their Completion Time is the longest (what is implicitly due to the use of temporal intention models). Combining the two emotional influences (E_{MI}-C) has also a long Completion Time, however, this condition resulted in a low Physical Interference Ratio and a Traveled Distance slightly shorter than the one of the NE-NC condition. Therefore, adding all the emotional processes of our approach to the group coordination (E_{MI}-C) makes it possible to reach a good compromise between security, distance, and speed while achieving the best Success Rate, demonstrating the usefulness of these emotional processes in improving the group adaptability. In dynamic and changing conditions, finding such compromise is probably the rational thing to do.

Future trends

In our future, work we intend to demonstrate the versatility and extensibility of the presented model in other settings by validating other emotional influences on an agent's decision-making mechanism in order to allow it to adapt its behavior to dynamically changing situations. Further extension of the model includes generation of positive emotions when, for instance, intentions are carried out by the agent's actions, when agent's actions satisfy multiple intentions, or when a temporary intention ends with a positive result. One main research avenue, which remains to be explored, is to make the agents able to modify their intention models through a learning mechanism, making them even more independent from knowledge provided by their designer. In order to widen the presented model's range of emotions and to better differentiate and refine elicited emotions, it could be beneficial to trigger an analysis of the current situation. Such an analysis could be contextualized from the agent's intention to make it richer than a simple continuous feature extraction from the agent's perceptions.

Models inspired by specific-stimulus theories of emotion will remain useful when improving reactions of artificial systems to specific situations. However, if complete autonomy of these systems is to be achieved, then provision must be made for a general self-analysis mechanism allowing them to detect situation for which their decision-making processes are not adapted. The main argument in support of this viewpoint is that adaptation of a decision-making mechanism to unforeseen situations is a costly process that cannot be executed continuously.

Conclusion

We believe situated agents' adaptability ultimately depends on the detection of the situations for which their regular decision-making process is not appropriate, calling for a behavioral or cognitive reaction. Detecting such situations is a key problem for situated agents because their environment is dynamic, continuous and unpredictable. Psychologists have identified that one of the functions of human emotion is to highlight these kinds of situations, allowing other cognitive processes to address them. We have developed an emotional process that allows situated agents to detect such situations by using temporal models of their intentions. This model does not rely on a priori knowledge of specific environmental conditions nor of the agent's mission objectives, but only on how control resources (i.e., behaviors) satisfy the agent's intentions. This independence between decision-making processes and human provided knowledge is crucial to an agent's autonomy since we cannot fully understand how an agent experiences its reality. Our long term objective is to provide situated agents with a generic self-analysis mechanism detecting situations requiring adaptation. Such a mechanism is required to autonomously trigger relevant learning phases and behavioral strategy switches and is therefore necessary to achieve complete autonomy for situated agents.

Acknowledgments

The authors gratefully acknowledge the contribution of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chair (CRC), in the support of this work.

References

Brooks, R. A. (1991). Intelligence without reason. *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91),* (pp. 569-595).

Brown, J. S., & Farber, I. E. (1951). Emotions conceptualized as intervening variables - with suggestions towards a theory of frustration. *Psychological Bulletin, 48*, 465-495.

Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Frijda, N. H. (1989). Les théories des émotions : Un bilan. In B. Rimé, & K. R. Scherer (Eds.), *Les émotions* (pp. 21-72). Paris: Delachaux et Niestle.

Gadanho, S. C. (2002). Emotional and cognitive adaptation in real environments. *Symposium ACE'2002 of the 16th European Meeting on Cybernetics and Systems Research.*

Goldberg, D., & Matarić, M. J. (2002). Design and evaluation of robust behavior-based controllers for distributed multi-robot collection tasks. In T. Balch, & L. E. Parker (Eds.), *Robot teams: From diversity to polymorphism,* (pp. 315-344), AK Peters.

Goldberg, D., & Matarić, M. J. (1997). Interference as a tool for designing and evaluating multi-robot controllers. *Proceedings of the National Conference on Artificial Intelligence (AAAI),* (pp. 637-642).

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Heise, D. (1979). *Understanding events: Affect and the construction of social action,* New York: Cambridge University Press.

Higgins, E. T. (1989). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319-340.

Lazarus, R. S. (1966). *Psychological stress and the coping process,* McGraw-Hill, New-York.

Lazarus, R. S. (1999). *Stress and emotion a new synthesis*. New York: Springer Publishing Company.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer Plublishing Company.

Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: W. W. Norton.

Mandler, G. (1997). *Human nature explored*. New York: Oxford University Press.

Michaud, F., Côté, C., Létourneau, D., Brosseau, Y., Valin, J., Beaudry, É, et al. (2007). Spartacus attending the 2005 AAAI conference. *Autonomous Robots, Special Issue on the AAAI Mobile Robot Competitions and Exhibition. 22*(4), 369-384.

Murphy, R. R., Lisetti, C. L., Tardif, R., Irish, L., & Gage, A. (2002). Emotion-based control of cooperating heterogeneous mobile robots. *IEEE Transactions on Robotics and Autotomation, 18*(5), 744-757.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*(4), 435-450.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1*(1), 29-50.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge, England ; New York: Cambridge University Press.

Parker, L. E. (1998). ALLIANCE: An architecture for fault tolerant multi-robot cooperation. *IEEE Transactions on Robotics and Automation, 2*(14), 220-240.

Piaget, J. (1989). Les relations entre l'intlligence et l'affectivité dans le développement de l'enfant. In B. Rimé, & K. R. Scherer (Eds.), *Les émotions.* (pp. 75-96). Paris: Delachaux et Niestle.

Picard, R. (1995). *Affective computing*. Cambridge, MA: MIT Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik, & H. Kellermann (Eds.), *Emotion; theory, research and experience.* (pp. 3-33), New York: Academic Press.

Rivard, F., Bisson, J., Michaud, F., & Létourneau, D. (2008). Ultrasonic relative positioning for multi-robot systems. *Proceedings of IEEE International Conference on Robotics and Automation.*

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). *Appraisal theories of emotions: Theories, methods, research,* New York: Oxford University Press.

Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition and Emotion, 7*(3), 325-355.

Scheutz, M. (2004). How to determine the utility of emotions. *Proceedings of AAAI Spring Symposium,*

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74*(1), 29-39.

Sloman, A. (2008). *Questions about emotions.* Retrieved 06-2008, from http://www.cs.bham.ac.uk/research/projects/cogaff/misc/emotions-questions.html

Robinson, D. T, SmithLovin, L., & Wisecup, A. (2006). Affect control theory. In J. E. Stets, & J. H. Turner (Eds.), *Handbook of the sociology of emotions* (pp. 179-202). New York: Springer.

Spinoza, B. (1677). *Ethique* [*Ethica ordine geometrico demonstrata*] (F. Alquié Trans.). (2nd ed.). Paris: Presses universitaires de France, 1966.

Stets, J. E. (2003). Emotions and sentiments. In J. Delamater (Ed.), *Handbook of social psychology* (pp 309-335) New York: Springer.

Vaughan, R. T., Gerkey, B. P., & Howard, A. (2003). On device abstractions for portable, reusable robot code. *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2421-2427).

Velásquez, J. D. (1998). A computational framework for emotion-based control. *Proceedings of Fifth International Conference on Simulation of Adaptive Behaviors (SAB'98),*

Watson, J. B. (1919). *Psychology, from the standpoint of a behaviorist*. Philadelphia: Lippincott.