A Hormonal Model of Emotions for Behavior Control

Dolores Cañamero
VUB Artificial Intelligence Laboratory
Pleinlaan 2, Room G–10–725
1050 Brussels, Belgium
lola@arti.vub.ac.be

Presented in the 4th European Conference on Artificial Life, ECAL'97, Brighton, UK, July 28–31, 1997

Abstract

This paper investigates some aspects of how emotions can affect the behavior of an artificial creature, a "newborn" endowed with motivational states and basic emotions. Different "agents" integrate this creature including agents for emotional phenomena. However, contrary to cognitive approaches, we do not tackle these directly, but go below the "agent level" and use physiological parameters to model them. Motivations drive behavior selection and organization based on the notions of arousal and satiation. Emotions exert further control by sending "hormones" that may affect the creature's perceptive, attentional, and motivational mechanisms, also modifying the intensity and execution of the selected behavior.

1 Introduction

Why would we need to introduce emotions in artificial autonomous agents? Emotions seem to have a central position in autonomy and adaptation in biological systems, and our claim is that we ought to introduce them in our designs if we want to build better adapted and more "life-like" creatures. This can mean several things—systems with better communication capabilities, with more flexible behavior, showing human/animal-like types of "errors", etc.—depending on the aspect of emotions we want to investigate. Different aspects have been stressed by diverse research paradigms [Cornelius 1996]. The *Darwinian* or evolutionary approach [Ekman 1992, Izard 1993] has focused on the study of the (facial) expressions and bodily responses accompanying emotions. These are adaptive—they represent our readiness to confront environmental contingencies encountered throughout our phylogenetic history—and therefore universal, at least a basic set of them. For the *Jamesian* tradition [Levenson *et al.* 1990], physiological changes are the cause of emotional experience. The *cognitive* approach [Ortony *et al.* 1988], focuses on the conscious aspects of emotions, which are considered to be based on appraisals. For *social constructivists* [Averill 1980] emotions can be fully understood at a social level of analysis, having nothing to do with remnants of our evolutionary past. *Neurobiology* [Damasio 1994, Vincent 1994, LeDoux 1996] attempts to understand emotions by investigating the fundamental mechanisms in the brain systems that give rise to them.

Our interest in emotions shares the same concerns as the neurobiology and evolutionary perspectives. The model we have implemented draws some inspiration from this research, within a "nouvelle AI" approach [Brooks 1991] that also takes Minsky's theory of intelligence [Minsky 1985] as a general framework. These three paradigms view intelligence as emerging from the interactions of "systems" connected in specific ways. Our creatures are composed of "agents", which includes agents for emotional phenomena as well. However, contrary to "cognitive" models of emotions in AI, we do not model emotions directly, at the "agent" level. We agree with neurobiology and "nouvelle AI" that to study complex emergent phenomena, one has to go below the level of the phenomenon we want to investigate [Pfeifer 1993]. We have therefore endowed our creatures with an artificial physiology that will allow us to ground them in the(ir) world and to act on their behavior at this level.

The paper is organized as follows. Section 2 examines some ideas that have inspired our design. Section 3 presents the microworld where our creatures (Section 4) live. Section 5 shows how emotions affect their behavior. Section 6 draws some conclusions and perspectives.

2 A Model of Emotions

The approach adopted here draws inspiration from the neurobiology and evolutionary traditions. From neurobiology, we adopt the following hypotheses:

- To understand emotions, we have to start with a study of their biological mechanisms—brain and hormonal systems—and how they evolved.

- Emotion and cognition are mediated by separate but interacting brain systems. The perceptual representation of an object and its "emotional meaning" are processed separately in the brain.

- Consciousness is not a necessary element to explain emotions. Emotional responses arise for the most

part in old evolutionary systems, and are therefore unconscious. As [LeDoux 1996] points out, when emotional systems function in animals that have capacity for conscious awareness, as humans do, conscious emotional feelings occur, but this cannot be claimed to happen in other animals.

Following [Kandel et al.1995], we distinguish three main components in the development of emotions: the recognition of an important event; an emotional experience in the cortex that mediates outgoing signals to peripheral structures; and reflexive autonomic and visceral responses. The causal relation between these components is not clear, though. We adopt Schachter-Damasio's "constructive" hypothesis [Schachter 1964, Damasio 1994] to see emotions as "stories" the brain concocts to explain bodily reactions.

From the evolutionary perspective, we adopt the following hypotheses:

- Emotions are biological functions of the nervous system involved in the survival of both the individual and the species. This does not imply that emotions cannot sometimes be dysfunctional and inadaptive on small time scales and at the individual level (why do we have so little control of our emotions?).

- The emotional brain systems being evolutionarily older than the cognitive ones, the affect system provides the primary blueprints for cognition and decision making, as well as for action.

- There seems to be a set of basic emotions, where "basic" has to be understood as both, being universal and being more closely related to survival.

- Each of these emotions may involve different brain systems that evolved for different reasons. Therefore, we implement the basic emotions as different agents that have specific behaviors and functions. "Complex" emotions—emotion blends and mixes—result from the combined effects of more basic emotions which are simultaneously active.

2.1 Emotions versus other Affective States

We distinguish between *motivations* and emotions, and assign them different roles in behavior selection and adaptation. Motivational states, such as hunger, thirst, aggression, etc. are drives that constitute urges to action based on internal bodily needs related to self-sufficiency and survival, whereas we view emotions as second-order modifiers or amplifiers of motivation. Neurobiology attempts to reduce the problem of motivation to that of a complex reflex regulated by excitatory and inhibitory control mechanisms in response to multiple stimuli, some

of them internal, some external. In general, motivations can be seen as homeostatic processes which maintain a controlled physiological variable within a certain range. Therefore, they involve arousal and satiation, and vary as a function of deprivation. According to [Kandel et al.1995], and in our implementation, motivations have three main functions: they steer behavior toward, or away from, a specific goal; they increase general alertness and energize the individual to action; and they combine individual behavioral components into a goal-oriented behavioral sequence. Motivations have to be integrated in artificial systems to promote decision making, activity selection, and autonomy. The mechanisms that put in relation the different components of emotional states are however less clear. Unlike motivations, the homeostatic model does not seem completely adequate in this case, in particular the notions of arousal and of activation as its indicator. Contrary to activation theories, which state a mere correlation between the amount of hormone released, amount of neural excitation, and amount of emotional arousal, we follow [Pribram 1984] to view emotional activation as an indicator of a change in configuration of neural and endocrine activity with respect to the habitual stable baseline of the organism. Another fundamental difference between motivational and emotional states is that, whereas motivations are highly urgent and specific, emotions combine both, structural/functional specificity and generality of object, time, intensity, and density [Tomkins 1984]. But if emotions show this generality, what accounts for the activation of different affects? Tomkins proposes three variants of a single principle: stimulation increase, which activates both positive and negative affects such as startle, fear, and interest; (high) stimulation level, which only activates negative emotions such as distress or anger; and stimulation decrease, which only activates positive emotions such as joy.

Emotions are also related to moods and temperaments. A *mood* can be explained as a (low) tonic level of arousal of an emotional system, whereas emotions imply sudden activation and short duration. Moods are not explicitly modeled in our implementation—we do not have agents for them—but low activation levels of emotions can be thought of as moods. *Temperaments* can be explained as "predetermined" threshold levels that make the activation of a particular emotion more likely. Temperaments have not been integrated in our system yet.

2.2 What are Emotions for?

Among the many possible functions that emotional states can have in adaptive systems, this work has focused on the following.

Emotions act as *amplifiers or modifiers of motivation* and behavior. Tomkins views emotions as the primary "innate" biological motivating mechanism, since in com-

plex systems, drives have insufficient strength as motives and need to be amplified by affects: "The affect system is the primary motivational system because without its amplification, nothing else matters, and with its amplification, anything else *can* matter. It thus combines urgency and generality" [Tomkins 1984, p. 164]. Accordingly, emotions seem to be of paramount importance in decision making [Damasio 1994].

Emotions and their expression are crucial in communication. An increasing number of research projects in AI and robotics are investigating these expressive aspects of emotions (see e.g. [Picard 1995, Velásquez 1996, Shibata *et al.* 1996]). Rather than on this "external" aspect, we have focused first on the role that emotions can play in contributing to maintain the internal stability of the organism, that we believe fundamental for adaptation. Although a homeostatic model is not enough to explain emotional states, emotions seem to play some role in homeostasis. [Pribram 1984] attributes to motivational and emotional mechanism complementary roles—while motivation is concerned with the operations of appetitive processes that try to activate action as a response to deprivation, emotion is derived from processes that try to stop ongoing behavior, i.e., it is concerned with satiety processes of re-equilibration.

Finally, emotions provide the primary blueprints for "cognition". Four main aspects attract our interest here: perception, attention, memory, and learning. Emotional states greatly influence perception and attention, both in level and object, and also promote selective memory and learning [Bower & Cohen 1982]. [Tomkins 1984] points out that the generality of time, object, intensity, and density characteristic of emotions are not the consequence of learning, but rather the structural innate features that make learning possible. Only the first two aspects are currently implemented in our creatures.

3   The Synthetic World

The microworld is a two-dimensional grid (Figure 1) inhabited by geometrical figures of three main types:

- Living beings, which fall into two species—Abbotts[1] and Enemies, both dot-shaped (occupying a single cell). Following [Brooks 1991], we call our autonomous agents "creatures".

- Food and water sources, also dot-shaped. They are randomly distributed in the grid when a scenario run starts. They can be consumed by both Abbotts and Enemies, at different rates depending on the intensity of the creature's needs. Once exhausted, food and water sources disappear and regenerate at a random location. Food can also be moved around by Abbotts.

- Inanimate blocks of diverse shapes and sizes—lines, triangles, squares, rectangles, and circles. Also randomly located when a new scenario starts, blocks cannot move by themselves. They serve two purposes, being obstacles in the navigation task, and "toys" that Abbotts can move around when playing.

All these entities are characterized by certain physical properties[2] which allow them to be perceived as objects by living beings:

Gravity level increasing downwards by one a row. The lower an object is located in the grid, the harder it is for a creature to move it.

Occupancy: empty (0), full (1), or partially full cell (2). Partially full cells correspond to the vertices and sides of triangles, and to the borders of circles.

Hardness of the object in the cell—an integer ranging from 0 (empty cell) to 4.

Brightness: an integer in the range 0–9.

Surface: list of cells occupied by an object[3].

Amount of organic matter: an integer in the range 0–5, where 0 corresponds to empty cells and inanimate blocks, and 5 to "healthy" Abbotts and Enemies[4], and to intact water and food sources. Any value in between corresponds to injured individuals or to partially consumed food or water. Living beings whose amount of organic matter goes down to 0 die.

This microworld has been implemented in CLOS, the object-oriented package on top of Common Lisp. All the entities in this world are thus objects, including the creatures and the agents that integrate them. Information about the physics of Gridland is represented in a 30 × 30 array of six-dimensional vectors, each containing numerical information about the features present in a cell.

4   Creatures

Gridland's population of living beings falls into two species—Abbotts and Enemies. Creatures consist of three types of elements: a set of *physical attributes* (introduced in Section 3), a set of *physiological variables* that define their bodily state, and a collection of *agents* of different types that can either send messages to other agents or switch them on and off. The design of our creatures responds to the following principles:

---

[1]Named after E.A. Abbott, author of the novel *Flatland*.

[2]Since space is discrete in Gridland, physical attributes are linked to each cell.

[3]This feature is not accessible to living beings, but used to moved things around in the animation.

[4]When a creature is bitten or attacked by another one, it loses a certain amount of organic matter that is dependent on the intensity with which it is attacked. Injured creatures can recover from this condition by eating.
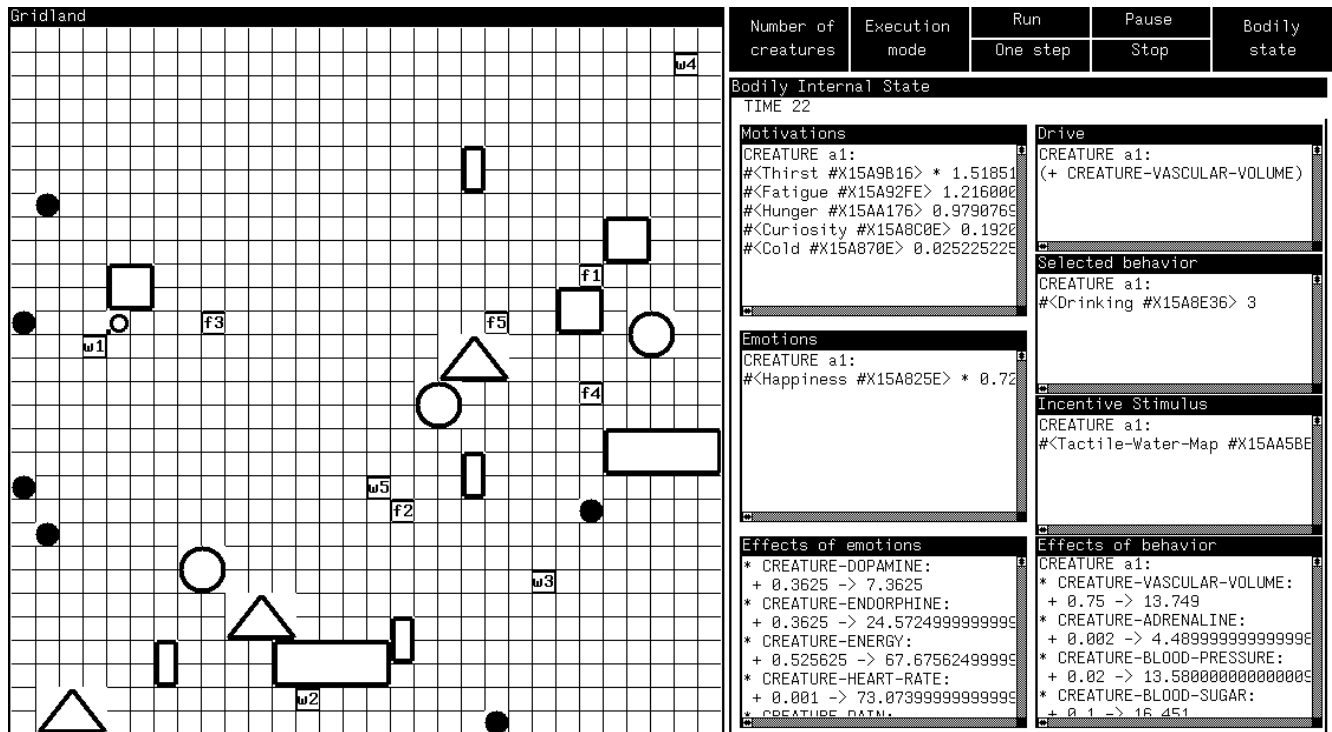
Gridland

| Number of creatures | Execution mode | Run | Pause | Bodily state |
| | | One step | Stop | |

Bodily Internal State
TIME 22

Motivations
CREATURE a1:
#<Thirst #X15A9B16> * 1.51851
#<Fatigue #X15A92FE> 1.216000
#<Hunger #X15AA176> 0.9790769
#<Curiosity #X15A8C0E> 0.1920
#<Cold #X15A870E> 0.025225225

Emotions
CREATURE a1:
#<Happiness #X15A825E> * 0.72

Effects of emotions
* CREATURE-DOPAMINE:
  + 0.3625 -> 7.3625
* CREATURE-ENDORPHINE:
  + 0.3625 -> 24.572499999999
* CREATURE-ENERGY:
  + 0.525625 -> 67.67562499999
* CREATURE-HEART-RATE:
  + 0.001 -> 73.07399999999999
* CREATURE PAIN:

Drive
CREATURE a1:
(+ CREATURE-VASCULAR-VOLUME)

Selected behavior
CREATURE a1:
#<Drinking #X15A8E36> 3

Incentive Stimulus
CREATURE a1:
#<Tactile-Water-Map #X15AA5BE

Effects of behavior
CREATURE a1:
* CREATURE-VASCULAR-VOLUME:
  + 0.75 -> 13.749
* CREATURE-ADRENALINE:
  + 0.002 -> 4.489999999999998
* CREATURE-BLOOD-PRESSURE:
  + 0.02 -> 13.580000000000005
* CREATURE-BLOOD-SUGAR:
  + 0.1 -> 16.451

Figure 1: Snapshot of Gridland. The empty "dot" with a black spot in one corner—its eye—is an Abbott. Enemies are shown as full dots.

- Complete creatures situated in a dynamic environment. The microworld is not a slice of the real world, but a different world with its own physics. The creatures embedded in it must sense "continuously" while acting, and they categorize it according to the structure of their perceptual system—no labeled percepts are provided. This sensory information will be processed by different agents at different levels and kept very distributed, as other agents will use it differently for different purposes.

- Developmental approach to intelligence, starting with a "newborn" creature that will have to grow. It is endowed with some mechanisms and structures to survive in its environment, such as a perceptual system embedding *a priori* notions of space and object, and motivational and emotional systems.

- Incremental design approach inspired from evolution [Brooks 1991]. To achieve more complex behaviors, new societies of agents have to be added without modifying already present structures.

- (Rough) inspiration from biology in different aspects, such as: the architecture of the creatures—the types of agents, their connections and communication patterns; the introduction of an internal bodily state that constitutes a synthetic autonomic nervous system and internal milieu; and the use of principles that

seem to govern several bodily mechanism, such as homeostasis and functional dualism [Vincent 1994].

4.1 Enemies

Enemies are a very simple type of creatures, with a rather stereotyped behavior, their main role being introducing more dynamicity in Gridland—they wander around continuously without hitting the objects they encounter, they try to ingest every form of organic matter[5] they run into, and they withdraw whenever they feel pain, caused either by other creature's attack or by their hitting a block. Their architecture includes only the elements necessary to produce this behavior. Enemies have nine agents: three sensors—an internal pain sensor and two 8-bit tactile sensors, one indicating the presence or absence of objects, and another tracking the amount of organic matter in each of their surrounding cells; one motor map—a 4-bit vector indicating the presence or absence of an object on each of their four motor directions, left, right, top, and bottom; two effectors—a foot and a mouth; and three behaviors: walking, eating, and withdrawing. Pain is the only physiological variable that takes part in the homeostasis, the amount of pain mapping the extent of the injury (i.e. the amount of matter taken away). Whenever pain is felt, a withdrawal behav-

---

[5] This includes food, water, and living beings, both Abbotts and other Enemies.

4

ior occurs in a direction different from the current motor direction. This behavior makes the pain disappear. The distance covered is proportional to the level of pain.

4.2 Abbotts

Abbotts show much more flexible behaviors made possible by a more sophisticated architecture. They serve as a testbed to investigate emotions.

4.2.1 Bodily state

Abbotts bodily state consists of a number of physiological variables that constitute their internal milieu. These variables can be of two types—"controlled variables" and "hormones".

Controlled variables are necessary for the creature's survival, and therefore must show a certain stability so that environmental changes do not put the creature's life in danger. They are controlled by homeostatic processes that keep them within a certain range. Abbotts controlled variables are: adrenaline, blood pressure, blood sugar, endorphines, energy, heart rate, pain, respiration rate, temperature, and vascular volume.

In this artificial internal milieu, control of these variables is achieved through the action of two mechanisms. The creature actions—the execution of a behavior—can directly modify the values of a subset of controlled variables that is specific to each behavior. The second mechanism is what we call *hormones*, which are released when Abbotts have an emotional reaction. Hormones can only act on specific receptors attached to the internal sensors that monitor the controlled variables. Each hormone can affect several variables, but it has a particular valence with respect to each of them; therefore, two different hormones are needed for each controlled variable. For example, release of insuline decreases the level of blood sugar, whereas glucagon increases it; endorphines decrease the level of pain, which is increased by substance P; vasopressin increases the vascular volume, angiotensin decreases it, etc. Hormones have two functions—communicative and synthesizing. On the one hand, they allow emotional states to selectively act on certain subsets of controlled variables—those responsive to them. On the other hand, they modify (increase or decrease, depending on the particular hormone) the amount of the controlled variables that are receptive to them. This way, emotional states contribute to keep the stability of the internal milieu necessary for the creatures survival and adaptation. Abbotts have also a second group of hormones that are not released by specific emotional states, but by general activation patterns, and that serve to control their level of arousal. Very high emotional states produce GABA, which causes Abbotts to enter in a "confused" perceptive and attentional state, whereas moderate emotional states release gluta-

mate that makes Abbotts to become more attentive and alert.

4.2.2 Agents

This section briefly presents the agents that make up Abbotts. A more detailed description and an account of how these agents relate to those in [Minsky 1985] can be found in [Cañamero 1997]. An agent is characterized by: name; owner—the creature to which it belongs; incentive stimulus to which the agent responds—a list of physical features, a physiological variable, or another agent; activation level—a real number; and state—1 if its stimulus has been perceived, 0 otherwise.

Sensory receptors can be of three kinds. *Internal sensors* monitor and convey interoceptive information concerning the variables defining Abbotts bodily state. They also include a list of hormones to which the variable they regulate can react. Two *somatic sensors* convey proprioceptive information: one tracks the position of the movable eye—one of the four corners of the body of Abbotts—the other the motor direction in body coordinates—top, left, bottom or right. *External sensors* provide information about the external world. They are of two types. *Tactile* sensors output a 9-bit vector reporting on the value of their respective incentive stimulus in each of the surrounding cells, plus the cell occupied by the Abbott. Abbotts have four tactile sensors: the gravity, occupancy, hardness, and organic amount sensors. Two *visual* sensors, one for brightness and another to measure the (square) distance (with respect to Abbotts single eye) of the occupied cells within a 5×5-cell, 90-degree visual field, output a list of integers.

Direction-nemes are agents associated with a particular direction or region in space. Abbotts have eight direction-nemes—one for each neighboring cell—corresponding to their different (sensory-motor) spatial regions in terms of body coordinates. Each has access to the information contained in a particular position of the output vector of one or more external sensors. They can be seen as sort of interneurons that play two main roles. First, they translate this information—an integer—into a format that can be understood by the other agents in the "brain"—a binary vector. Second, they perform some sensory integration in the tactile pathway, as information concerning occupancy and hardness are assembled in a single vector. The only information they take into account in the visual pathway is brightness. Both pathways are kept separate. Direction-nemes also communicate with agents devoted to higher-level sensory processing—object recognizers. They send their stored patterns to the recognizers, which return encoded

information[6] about the object category to which the pattern corresponds; however, this output is meaningless to direction-nemes. The effect of these agents is some sort of Kantian *a priori*—Abbotts can only perceive and act within a spatial framework.

Recognizers are higher-level processors of complex sensory patterns. Only object recognizers have been used so far, one for the tactile pathway and another for the visual one. They implement ART–1 neural nets [Carpenter & Grossberg 1988] that perform vector clustering for unsupervised, competitive learning. Recognizers take as input binary vectors and perform a match-based learning process that results in either the pattern being recognized under a known category, or a new category being formed and assigned to an empty output unit. They output the index of the winning unit. Some features of ART–1 nets make them particularly suited for their use in dynamic environments: automatic switch between learning and recognition modes, one-shot learning/recognition (with fast learning), and the existence of a parameter—the vigilance threshold—that allows to play with the granularity level at which objects will be recognized and learned[7]. Recognizers incorporate two additional mechanisms: deletion of the less frequently used categories, and reinitialization of the network's weights and modification of the vigilance threshold—and of the granularity of the recognition or categorization—depending on the creature's emotional state.

Maps are agents that represent the presence of high-level sensory information in a topology-preserving way. They are fed by direction-nemes—all maps share the same direction-nemes, but each map is reactive to a single incentive stimulus—a binary feature vector—and is hence specialized to represent only one type of information. Since maps act as incentive stimulus for different motivations and behaviors, we don't want a general map that centralizes all the high-level sensory information. Abbotts have two main categories of maps—tactile and visual. Maps contain a binary vector with a position for each direction-neme they communicate with, which indicates the presence or absence of their incentive stimulus at that location. To obtain this information, each map sends its incentive stimulus to the corresponding (tactile or visual) recognizer as training signal, and the recognizer outputs the index of the winning output unit. The map then compares this index with the corresponding one stored by each direction-neme. If the index is the same, the stimulus has been observed at that particular location and the concerned position is set to one. Maps

---

[6]Object recognizers implement competitive neural nets, and return the index of the winning neuron for that pattern.

[7]This somehow amounts to endowing the agent with an *a priori* perceptual structure that determines the kinds of objects our creatures will be able to know.

Behavior	Stimulus	Main Effect
Attack	living-being	− adrenaline
Drink	water	+ vascular volume
Eat	food	+ blood sugar
Play	block, Abbott	+ endorphine
Rest	top of flat block	+ energy
Walk	free space	+ temperature
Withdraw	pain	− pain

Table 1: Abbotts behaviors.

activate themselves if the stimulus has been observed at least at one location. Abbotts have five tactile maps—occupancy, water, food, living-being, and block maps—and six visual maps—the same as above except that the living-being-map is split into a visual-enemy-map and a visual-abbott-map.

Effectors are agents that perform motor actions in the world, with given intensities and directions. Abbotts have three effectors: a hand, a foot, and a mouth. Each can perform a motor action in response to another agent's activation—a behavior or an affective state. Hand's motor actions are open, close, push, and pull. Foot's motor actions are go-up, go-left, go-down, go-right (all with varying intensities), and stop. The mouth's only action is ingest.

Behaviors correspond to consumatory behaviors in the ethology literature—those contributing to the balance of resources that ensure a creature's self-sufficiency. They implement goal-achieving systems [McFarland 1995] that can recognize that a stimulus or goal is present when they encounter it—preprogrammed recognition is thus their main feature—and change their behavior accordingly, but the process of arriving at the goal is determined by the environment. A behavior agent is defined by: a preprogrammed incentive stimulus—a map or a sensor; a certain intensity and direction determined by both, environmental and internal factors; an effector; and a list of effects (increase or decrease) that the execution of the behavior has on the creature's physiological variables at every time step, the first of which is the main effect of the behavior. The extent to which the execution of a behavior affects the values of these variables is dependent on the behavior and proportional to its intensity. A behavior can be executed only if it has been selected by the motivational/emotional state of the creature, and its incentive stimulus is being observed, i.e., the corresponding agent (map or sensor) is active. When this second condition does not hold, the motivational system will either look for another behavior that can satisfy the current need, or call a manager agent. Abbotts current behaviors, their incentive stimuli and their main effects are shown in Table 1.

Motivation	Drive
Aggression	decrease adrenaline
Cold	increase temperature
Curiosity	increase endorphine
Fatigue	increase energy
Hunger	increase blood sugar
Self-protection	decrease pain
Thirst	increase vascular volume
Warmth	decrease temperature

Table 2: Abbotts motivations and their drives.

Managers implement appetitive behaviors: those that make more likely that the conditions that allow to satisfy a basic need (the presence of a stimulus) hold. Abbotts current managers are the finder, look-for, and go-toward agents. Unlike behaviors, they do not have pre-programmed recognition, but obey to what Minsky calls the *exploitation principle*—they make use of the activity of other agents without understanding how they work. Contrary to the appetitive behaviors in [Maes 1991], which are stimulus-specific, managers can respond to any stimulus[8] that another agent "tells them" to attend to. This stimulus has the form of one or several agents—maps or sensors—that the manager will try to turn into an active state. Therefore, managers exhibit a goal-directed behavior [McFarland 1995] guided by an explicit representation of the goal to be achieved. Managers are exploited by motivations, and they can exploit other managers, behaviors, or simple motor actions.

Motivations implement a homeostatic model. They are characterized by: a controlled (physiological) variable; an external incentive stimulus that can increase the motivation's activation level, but cannot trigger it; an error signal or drive—increasing or decreasing the level of the controlled variable; and a satiation criterion. Table 2 shows Abbotts motivations with their corresponding drives. Each motivation receiving an error signal from its feedback detector receives an activation level proportional to the magnitude of the error, and an intensity calculated on the basis of its activation level. Several motivations can be active at the same time, but that with the highest activation level gets the creature's attention and tries to organize its behavior so as to satisfy its drive.

Emotions are agents that amplify or modify the motivational state of a creature and its perceived bodily state. They are characterized by: an incentive stimulus; an intensity proportional to the level of activation; a list of

---

[8]This use of the exploitation principle allows us to avoid having to program a different behavior for each potential incentive stimulus (e.g., go-toward-food, go-toward-water, etc.), without however making use of classical variables.

Emotion	Triggering event
Fear	Enemy
Anger	accomplishment of a goal menaced or undone
Happiness	achievement of a goal; Abbott
Sadness	inability to achieve a goal
Boredom	repetitive activity
Interest	block; Abbott; novel object or event

Table 3: Innate external stimuli triggering emotions.

hormones released when active; and a list of physiological manifestations. Due to their double aspect of specific and general mechanisms, emotions cannot be activated only by innate (internal or external) stimuli. They are activated and discriminated by:

- External events, either an object or the outcome of a behavior. These events can be prewired (Table 3), as a remnant of Abbotts phylogenetic past, or memorized (not yet in Abbotts). The perception of these "emotionally significant" events is preattentive, and therefore different—processed separately—from the "cognitive" perception of the same objects.

- General patterns of stimulation—sudden increase, high level, sudden decrease. This general activation principle makes that the same emotion can be felt in different circumstances. For instance, a sustained abnormal high level of any variable activates the anger agent. This way, emotions help control the homeostatic processes characterizing motivations.

- Particular patterns of physiological variable values specific to each emotion, which allow to distinguish between emotions activated by the same general mechanism, such as fear (high heart rate and low temperature) versus interest (low heart rate).

Several emotions can be active at the same time, and all of them release "hormones" with varying levels of intensity depending on the activation level. Emotions are activated by applying the above principles in the following order. An external event is the strongest one, and can activate an emotion by itself. If several or no events are observed, the second one is used for emotion selection and for the assignment of the corresponding activation levels. Since the second principle often selects more than one possible emotion, the third one is used to discriminate.

4.3 The Sensing-Acting Loop

Gridland time is discrete, proceeding in cycles. Scenarios can be run in two modes: automatic, for a number of steps set by the user, or stepwise. The user can also stop or interrupt a run to examine the state of the agents inside the creatures. The following steps are repeated at every clock tick:

1. The activation level of all agents and their activation state are (re)set to 0.

2. Both the internal variables and the environment are sensed, objects "subliminally" recognized, maps built. Not all this information will be attended to by Abbotts, but only those pieces that are relevant to their motivational state.

3. The effects of the creature's emotional state are computed and motivations assessed. The motivation with the highest activation is selected.

4. The active motivation selects the behavior(s) that can best satisfy its drive—consumatory if the incentive stimulus is present, appetitive otherwise.

5. The behavior is executed by the corresponding effector, which modifies both, the conditions of the world and the creature bodily state.

The reason for reinitializing the activation state and level of all the agents at every cycle (step 1) is that it is through modification of the physiological variables composing the creature's bodily state that the agents affect the behavior of the creature. We only keep a memory of the past state at the physiological level, the agent level being "recomputed" at every cycle. For the same reason, the user cannot modify the creatures agents, but only reset their physiological variables. Let us now examine how motivations and emotions contribute to the behavior of the creatures.

5 Using Emotions for Behavior Control

5.1 Behavior Selection

Behavior selection is accomplished by the motivation that gets the creature's attention. First, it looks for the behavior(s) that can best contribute to its satisfaction—those whose main effect coincides with the drive. If none is found, it will select a list of behaviors that can contribute to it to a lesser extent. If more than one behavior is selected, all of them are kept to try to opportunistically execute the first whose incentive stimulus is observed. If the behavior(s) cannot be executed, its sends a failure message back. Then the motivation activates the finder agent, providing it with its own stimulus—the agent that the finder must try to turn active—and the intensity of the urge, which will be transmitted by finder to the agents it switches on—other managers, behaviors, or motor actions. In the case the behavior becomes active, it is executed with a given intensity—that of the urge—which has an impact on the values of some physiological variables. The way in which the behavior contributes to the satiation of the drive (and to the modification of other variables) depends on the intensity with which it is executed. For some behaviors, such as withdraw

or attack, the intensity determines the strength of the motor actions; for others, the duration of the behavior—how many time steps it will last, provided that no external event (e.g., an Enemy's bite) or internal bodily state make other motivation more urgent. After an interruption, a creature can return to its previous motivation if it is still strong enough.

5.2 Adding Emotions

Emotion agents have several effects in Abbotts, namely in their perceptual, attentional, motivational, and behavioral mechanisms. Emotions act fundamentally by releasing hormones that alter the physiology of the creature (and its perception of it). The release of hormones affects selectively those agents that have receptors which are reactive to each particular hormone. This has two main effects, both proportional to the intensity of the emotion: a modification of the reading of the internal sensors, and a much smaller modification of the real value of the corresponding variable, which will only be felt at the next time step. This double effect wants to reflect the fact that the visceral effects of emotions are usually much slower than the behavioral reaction.

Emotions can modify Abbotts perception of both, their own bodily state and the external world. The former happens through hormonal modification of the internal sensor reading (e.g., in an euphoric emotional state, the happiness agent releases endorphine, which reduces the perception of pain). The latter is achieved by the modification of the vigilance threshold in recognizer agents, which has an impact on the granularity of the categories formed. A very high activation level of any emotion causes a "confused" state that lowers the threshold, producing a coarser object categorization, whereas a moderately high activation level causes an "alert" state that raises the threshold, producing a finer categorization. These modifications have also an impact on the performance of the creatures, as follows. The different maps can get "wrongly" activated as a consequence of an "abnormal" categorization; for example, a tactile block map can become active in the presence of an Enemy when the Abbott is "confused". Since map agents are used as incentive stimulus for behaviors, the fact of the tactile block map being active will activate the resting behavior if the creature is tired (i.e. the fatigue motivation is active); the Abbott will then try to sleep on top of the Enemy, be bitten by the it, and feel pain.

Emotions also modify motivations. The effects of emotions (i.e. of the hormones released) are computed before motivations are assessed, producing a modification of the creature's motivational state that can be seen as a sort of "inhibition" mechanism. The error signals that the motivation agents have can then be different from those they would have had in the absence of emotions, and therefore their activation levels and intensity will be dif-

ferent as well. This can either change the priority of motivations, which modifies what the creature attends to and what behavior gets selected, or change the way in which the behavior is executed in the case of a modification of the motivation's intensity, since this intensity is passed to behaviors. In extreme cases, this can prevent the execution of a behavior if its intensity is too low. The particular hormones released by the different emotions also affect the ways the behavior is executed. For example, an angry creature executes motor actions with more strength, as acetylcholine is released, whereas a sad creature will behave at a very slow pace under the inhibitory effects of GABA.

5.3 Discussion

Endowing our creatures with motivations makes them self-sufficient and autonomous, fit to survive in their dynamic environment, to make rapid choices that lead to the satisfaction of their needs and to take advantage of what the world offers to them, or to look for it if needed. They show both, goal-oriented and opportunistic behaviors; they are able to keep working on a task long enough to get a need satisfied, without neglecting what is going on in the world (a hungry Abbott looking for food will stop to drink if it begins to be also thirsty and water has been detected), or other more urgent needs that could arise (it will stop an activity to avoid an Enemy's bite if this latter is considered more important).

The addition of emotions in our creatures allows them to show some interesting phenomena, and in particular a much more flexible behavior. For example, the presence of an Enemy will activate the fear emotion, as it is an incentive stimulus prewired in Abbotts emotional system. However, this emotion does not determine a unique behavior in the creature. It releases both, tyroxin, which increases the level of adrenaline, and substance P, which increases that of pain. Depending on the levels of those two variables, the subsequent motivation of the creature can be either aggression, which leads to an attacking behavior, or self-protection, which produces withdrawal.

This example let us observe another interesting emotional phenomenon: the chaining of typical emotion sequences. Strong fear will be followed by anger, as it causes a high stimulation level of a controlled variable—adrenaline or pain. Depending on the motivation that became active—aggression or self-protection—this will lead to an amplification of the aggression motivation and of the intensity of the attack, or to a fear-withdraw-anger-attack chain. Moreover, a happy emotion or mood, depending on the intensity of the emotion previously felt, will follow in both cases, due to a decrease in the level of the controlled variable.

We have already seen in Section 5.2 how emotions influence perception and attention through a general activation pattern. Specific emotions can have particular effects on them, too. For instance, boredom releases GABA, which lowers the level of attention.

Let us consider a last example in which emotions clearly help maintain the stability of the internal milieu. Emotions such as sadness and boredom release the inhibiting hormone GABA that, besides the level of attention, also decreases the energy level. Since this level is taken into account to calculate the strength of motor actions, sad or bored creatures behave at a slower pace, consume less energy, and therefore contribute to save this precious resource necessary for the creature's survival.

6 Conclusions and Future Work

In this paper, we have investigated some aspects in which emotions can affect the behavior of an artificial creature inhabiting a 2-dimensional world, a "newborn" endowed with motivational states and a set of basic emotions. In particular, we have seen how emotions serve as amplifiers or modifiers of motives, allowing for more flexible behaviors than those shown by a creature driven only by motivations, and how they affect attention and perception, both of the own body and of the outside world. Contrary to cognitive approaches, we have modeled both concepts at the physiological level. Motivations constitute the basic mechanism that regulates the stability of the creature's internal milieu, necessary for survival, adaptation, and autonomy, driving behavior selection. They can be seen as implementing an "innate" basic value system. Emotions further contribute to maintain this stability, acting as a second-order mechanism on top of motives. Although we have adopted the hypothesis of the existence of a set of basic emotions modeled as agents, the use of an artificial physiology and internal milieu allows for other behaviors and emotional phenomena, such as the formation of typical emotion chains, to emerge.

We believe these results provide a first partial answer to the question we posed at the beginning of the introduction. Many more aspects of emotions can be investigated, though. This will of course imply adding more machinery to the creatures. In particular, their effects in memory and learning mechanisms attract our interest as the next research topic. Emotions seem to offer many possibilities here—determining what is worth remembering and learning, taking advantage of the emotional state in which something important happened for selective memorization and retrieval, using emotional states for reward and punishment, etc. An interesting issue is learning of emotions themselves—assigning valences to emotions and to association of emotions and events, so that the creature can learn to somehow control (promote or inhibit) its emotions; e.g., it may "discover" that fighting an Enemy is more useful than fleeing away under certain conditions, and modify its internal bodily state so as to make this possible.

Acknowledgements

References

[Averill 1980] Averill, J.R. 1980. A Constructivist View of Emotion. In R. Plutchik and H. Kellerman, eds. *Emotion: Theory, Research and Experience*, Vol. 1, 305–339. New York, NY: Academic Press.

[Bower & Cohen 1982] Bower, G.H. and Cohen, P.R. 1982. Emotional Influences in Memory and Thinking: Data and Theory. In M.S. Clark and S.T. Fiske, eds. *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition* 291–331. Hillsdale, NJ: Erlbaum.

[Brooks 1991] Brooks, R.A. 1991. Intelligence Without Representation. *Artificial Intelligence* 47(2): 139–159.

[Cañamero 1997] Cañamero, D. 1997. Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In W. Lewis Johnson, ed., *Proceedings of the First International Conference on Autonomous Agents*, 148–155. New York, NY: ACM Press.

[Carpenter & Grossberg 1988]
Carpenter, G.A. and Grossberg, S. 1988. The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network. *Computer*, March: 77–88.

[Cornelius 1996] Cornelius, R.H. 1996. *The Science of Emotion*. Upper Saddle River, NJ: Simon & Schuster.

[Damasio 1994] Damasio, A.R. 1994. *Descartes' Error*. New York, NY: G.P. Putnam's Sons.

[Ekman 1992] Ekman, P. 1992. An argument for Basic Emotions, *Cognition and Emotion* 6(3/4): 169–200.

[Izard 1993] Izard, C.E. 1993. Four Systems for Emotion Activation: Cognitive and Noncognitive Processes, *Psychological Review* 100(1): 68–90.

[Kandel et al. 1995] Kandel, E.R., Schwartz, J.H., Jessell, T.M. 1995. *Essentials of Neural Science and Behavior*. Norwalk, CT: Appleton & Lange.

[LeDoux 1996] LeDoux, J. 1996. *The Emotional Brain*. New York, NY: Simon & Schuster.

[Levenson *et al.* 1990] Levenson, R.W., Ekman, P., Friesen, W.V. 1990. Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity, *Psychophysiology* 27: 363–384.

[Maes 1991] Maes, P. 1991. A Bottom-Up Mechanism for Behavior Selection in an Artificial Creature. In Meyer, J.-A. & Wilson, S.W., eds. *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, 238–246. Cambridge, MA: The MIT Press.

[McFarland 1995] McFarland, D. 1995. Opportunity versus Goals in Robots, Animals and People. In Roitblat, H.L., Meyer, J.-A. eds. *Comparative Approaches to Cognitive Science*. 415–433. Cambridge, MA: The MIT Press.

[Minsky 1985] Minsky, M. 1985. *The Society of Mind*. New York, NY: Simon & Schuster.

[Ortony *et al.* 1988] Ortony, A., Clore, G.L., Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge, UK: Cambridge University Press.

[Pfeifer 1993] Pfeifer, R. 1993. Studying Emotions: Fungus Eaters. *Proceedings of the European Conference on Artificial Life, ECAL'93*, 916–927. ULB, Brussels, Belgium, May 24–26.

[Picard 1995] Picard, R. 1995. Affective Computing. MIT Media Laboratory Technical Report No. 321. Cambridge, MA.

[Pribram 1984] Pribram, K,H. 1984. Emotion: A Neurobehavioral Analysis. In Scherer, K.R. & Ekman, P. *Approaches to Emotion*, 13–38. Hillsdale, NJ: Lawrence Erlbaum.

[Schachter 1964] Schachter, S. 1964. The Interaction of Cognitive and Physiological Determinants of Emotional States. In Berkowitz, L. ed. *Advances in Experimental Social Psychology* Vol. 1, 49–80. New York: Academic Press.

[Shibata *et al.* 1996] Shibata, T., Inoue, K., Irie, R. 1996. Emotional Robot for Intelligent System: Artificial Emotional Creature Project. *5th IEEE International Workshop on Robot and Human Communication*, 466–471. IEEE Press.

[Tomkins 1984] Tomkins. S.S. 1984. Affect Theory. In Scherer, K.R. & Ekman, P. *Approaches to Emotion*, 163–195. Hillsdale, NJ: Lawrence Erlbaum Associates.

[Velásquez 1996] Velásquez, J. 1996. *Cathexis: A Computational Model for the Generation of Emotions and their Influence in the Behavior of Autonomous Agents*. MIT Media Laboratory MSc Thesis. Cambridge, MA.

[Vincent 1994] Vincent, J.-D.. 1994. *Biologie des passions*. Paris: Editions Odile Jacob.