FAKULTÄT FÜR !NFORMATIK

Psychoanalytical Defense Mechanisms Applied to Autonomous Agents

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Wirtschaftsingenieurwesen Informatik

eingereicht von

Christiane Riediger
Matrikelnummer 0201162

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung:
Betreuer: O. Univ. Prof. Dipl.-Ing. Dr.techn. Dietmar Dietrich
Mitwirkung: Dipl.-Ing. (FH) Roland Lang

Wien, 10.03.2009 _____ _____
(Unterschrift Verfasser) (Unterschrift Betreuer)

Kurzfassung

Durch den Einsatz von Technik wird die Arbeit von Menschen unterstützt und erleichtert. Der Drang nach Robotern, die diese Arbeiten selbstständig durchführen, wir immer größer. Die Herausforderung dabei an autonome, mobile Roboter besteht darin, neue Situationen zu interpretieren und angemessen in unbekannten Umgebungen zu reagieren. Dabei ist es notwendig, dass der Roboter seine Umgebung wahrnehmen und aufgrund von bereits erlernten Situationen neue Entscheidungen treffen kann. Für die Entscheidungsfindungseinheit gibt es verschiedene Ansätze.
Ein Ansatz, für den sich das ARS Team am Institut für Computertechnik auf der technischen Universität Wien entschieden hat, ist die Umsetzung psychoanalytischer Modelle in die Technik. In diesem Modell unterscheidet Freud zwischen den drei Instanzen *Ich*, *Es* und *Überich*. Das *Es* ist jene Instanz, die für menschlichen Triebe zuständig und seit unserer Geburt vorhanden ist. Das *Überich* ist jene Instanz, die Gebote und Verbote beinhaltet, und sich im Laufe der Kindheit entwickelt. Das *Ich* ist jene Instanz, die die Verbindung zur Außenwelt darstellt und aufgrund von äußerer Wahrnehmung und Verboten bzw. Geboten des *Überich* die Triebe und Wünsche des *Es* versucht zu befriedigen. Des weiteren verwendet Freud in diesem Modell Konzepte wie Emotionen und Gefühle, die ebenfalls die Entscheidungsfindung beeinflussen.
Die Entscheidungsfindung, welche von Trieben, Wünschen, Emotionen und Gefühlen beeinflusst wird, bringt jedoch auch einige Probleme mit sich. Ein Trieb, der aufgrund von äußeren Gegebenheiten nicht befriedigt werden kann, wird Emotionen negativ beeinflussen. Weiters können dadurch Aktionen blockiert werden, die es ermöglichen würden den Trieb zu einem späteren Zeitpunkt zu befrieden. Als eine mögliche Lösung für dieses Problem ist der Einsatz von psychoanalytischen Abwehrmechanismen, die das Thema dieser Arbeit darstellen. Diese sind ebenfalls von Sigmund Freud entdeckt und erforscht worden. Seine Tochter Anna Freud hat sich nach seinem Tod damit weiter beschäftigt und seine Theorie erweitert. Um den Einsatz von psychoanalytischen Abwehrmechanismen für autonome Agenten zu ermöglichen, wurden die unterschiedlichen Mechanismen analysiert und ein allgemeines technische Modell entwickelt. Anschließend wurden vier Abwehrmechanismen technisch umgesetzt und in einem Simulator, der in Java programmiert wurden, getestet. Die Resultate der Simulation zeigen, dass die Abwehrmechanismen es ermöglichen, negative Emotionen wie Wut leichter zu kontrollieren und die Wiederherstellung der innere Balance des Agenten erleichtern.

Abstract

The human work is supported and facilitated by using technics. But the desire to have robots, which act alone, becomes more stronger. The challenge for autonomous, mobile robots is that it has the possibility to interpret new situations and to react in unknown situations appropriated. Therefor it is necessary that the robot can perceive the environment in which it should act and can make decisions due to the earlier learned situations. For the decision unit different approaches exist.

The approach, which is used by the ARS team of the Institute of Computer Technology on the Vienna University of Techology, is technical realization of psychoanalytical models. In this model he differed between the three instances *Ego*, *Id* and *Superego*. The *Id* is responsible for the human drives and exists since birth. The *Superego* contains the forbiddances as well as the restrictions and demands of the parents and evolves during the childhood. The *Ego* is the connection to the outer world and tries to satisfy the drives of the *Id* by including the conditions of the outer world as well as the forbiddances and restrictions of the *Superego* in the decision making process. Furthermore Freud used in his model concepts like emotions and feeling, which also influences the decision making process.

Decision making, which is influenced by drives, desires, emotions and feelings, implicates some other problems. A drive, which can not be satisfied, can influences emotions negative. As well this drive can block actions, which make it possible to satisfy it later. One possible solution for this problem is using psychoanalytical defense mechanisms, which are the topic of this work. These mechanisms are found and studied by Sigmund Freud. His daughter Anna Freud had resumed this research after his death. To enable the usage of the defense mechanisms in autonomous agents, various mechanisms are analyzed and a general technical model was developed. Afterwards four defense mechanisms are technical realized and tested in a simulator, which was programmed in Java. The results of the test are that defense mechanisms enables the agent to control easier negative emotions like rage as well as facilitates the rebuilding of the inner balance.

Acknowledgments

First of all I would like to thank Prof. Dietmar Dietrich for the opportunity to write this thesis at the Institute of Computer Technology. Special thanks go to my supervisor Roland Lang for the great support to my thesis. Further, I would like to thank Brit Müller and Anna Tmej, our psychoanalysts, for fruitful hints and discussions. My deep gratitude goes to my parents, for the opportunity to study and their support during my student's career. Furthermore, I would like to thank my boyfriend Bogumil for loving and supporting me during my student's career.

Table of Contents

Abbreviations

AI	Artificial Intelligence
ARS	Artificial Recognition System
BFG	Bubble Family Game
CI	Cognitive Intelligence
HTML	Hypertext Markup Language
ID	Identifier
XML	Extensible Markup Language

1 Introduction

To support and facilitate human work, technics is used. The desire to have robots, which work alone, becomes stronger. But to enable an agent to act without help in an unknown environment, various challenges like the high number of data points of the sensors have to be handled. This task can roughly be devided into three parts - the part 'How the agent perceives the environment interacting with', the part 'Decision making' and the part 'How the agent can interact with its environment'. In this work the part 'Decision making' is the interesting part. Various approaches try to solve the problem, how an agent can make decisions and interact in an unknown environment without hurting itself as well as human. Nevertheless the task should be done in an efficient way.

Researchers of the communities of Artificial Intelligences (AI) and Cognitive Intelligence (CI) have started an approach, which is divided in various phases. The first step was the adaption of the psychological principle of symbolization. The result of this step is the knowledge representation. Afterwards statistical methods and learning algorithms were applied. But it was realized that the human mind depends on an individual body, whereas it never minds if it is virtual or not. Therefore engineers were concentrating on the term 'embodiment' during the third phase. As well engineers have the goal to implement a representation of the outer world in the agent and the agent only operates relying on this internal knowledge base. The next phase can be described by searching for definitions of emotions and the corresponding feeling. This is done to find a way for implementing them into an agent [DFKU09]. Currently the fifth phase has started. The ARS (Artificial Recognition System) team on the Institute of Computer Technology of the Vienna University of Technology has started a project, which is inspired by the second topographical model of S. Freud. A technical realization of the psychoanalytical model is the aim of the team by following a bionic approach.

The ARS team has developed two model of the human mind. The first model has been designed by using knowledge of various sciences like psychology, neurology and psychoanalysis. But it has been a design flaw because of mixing different theories, which are sometimes incompatible. That is the reason why the second model is only based on one theory namely on the second topographical model of S. Freud. But to use this model for a technical implementation, the terms using by psychoanalysis have to be translate into a technical language. This is done by a team, which consists of engineers and psychoanalysts. In the second topographical model Freud distinguished between the three instances *Ego*, *Id* and *Superego*. The *Id* is responsible for the human drives and exists since birth. The *Superego* contains the forbiddances as well as the restrictions and demands of the parents and evolves during the childhood. The *Ego* is the connection to the outer world and tries to satisfy the drives of the *Id* by including the conditions

of the outer world as well as the forbiddances and restrictions of the *Superego* in the decision making process. Furthermore Freud used in his model concepts like emotions and feeling, which also influences the decision making process. But decisions, which are influenced by drives, desires, emotions and feeling, lead to other problems. If e.g. a drive can not be satisfied, but has high priority, may block other actions with lower priority. So the agent would like to fulfill a drive, which is not possible at that time. It might be possible that after some other actions the drive can be satisfied. So it is necessary to be able to reduce the priority of the not fulfilled drive. A solution for such a problem are psychoanalytical defense mechanisms. These mechanisms had been researched by S. Freud. His daughter A. Freud resumed the researches and developed them further. The defense mechanisms are the topic of this work and are analyzed for being able to use them in the decision unit of an autonomous agent.

In the first models of solving the problem that an agent should do its task more efficiently, was the application of emotions. As explained above, emotions are a part of Freuds research. But systems, which use emotions, differ significantly concerning to their aims and assumptions. In [PB09] a small number of projects, in which emotions are implemented, are explained in detail. It is seen that emotions improves the efficiency, if the agent was rewarded when it does its work because positive emotions like joy rises. Because of being able to set only one action after the other, the agent needs the possibility to decide which action has more priority. At that point the problem starts. The agent has to have to possibility to decide which action should be done and what should happen with the other action. In Freuds model drives, which are the signal of the needs of the body to the mind, influences our behavior as well as emotions do. So the decision which action should be set is influenced by drives and emotions. If a drive becomes not satisfied, it leads to frustration. Such a situation should be avoided. A solution therefor is the use of defense mechanisms. The advantage of using defense mechanisms is that various defense mechanisms are defined. The use of them depends on the conflict which occurs in the decision making unit. So depending on the kind of conflict, a defense mechanism can be selected for solving the problem. As example the problem explained in the paragraph before - a not fulfilled desire blocks other actions - can be solved e.g. by the defense mechanism Repression which means that the agent does not know the drive. Another possible solution is that the drive is sublimated by another drive which can be satisfied. This mechanism is called Sublimation. So the defense mechanisms enables the agent to solve the same problem in different ways. This is an advantage because the solution for a problem can be adjusted on the conditions of the outer world.

A. Buller decides to design his decision unit by using psychodynamics and not psychoanalysis. For him, the research of Freud includes controversial statements about sexuality and aggression. So he only uses a severely expurgated version of Freudian work which are psychodynamic concepts. He argues that AI can benefit from the concepts of Freud [Bul05]. In further work he expands its robots with a controller named Volitron. This controller is also arranged with defense mechanisms, which becomes activated if the roboter can not fulfill satisfactory plans for a too long span of time.

Defense mechanisms are a function of the instance *Ego*. They enable the *Ego* to decide which thoughts of the mind becomes aware. So the *Ego* has the possibility to distinguish between allowed thoughts and forbidden stimuli. If the *Ego* does not allow the satisfaction of a drive, the drive can be repressed or distorted or the instinctual aim is displaced, suppressed or inhibited. Specific defense mechanisms are evolved by the human mind [DFZB08]. To be able to use the mechanisms which are defined in [Fre84] or in [SSK97], they are analyzed and a general model was designed. Afterwards this model is transformed into a technical model. The implementation of four defense mechanisms is done by using XML files and Java. Afterwards the tests are carried

out in two different test benches - the functionality test bench and the prioritization test bench - as well as a simulation is done. The simulator therefor is the Bubble Family Game (BFG), which was developed by the ARS team and is the first prototype of the first model of the project ARS.

In this chapter a short overview about the area of research is given first. Afterwards the different phases of the research are presented. Then the project ARS is explained in which this work is embedded. The motivation for the work and some advantages of the chosen solution are presented. Another approach, which is done by A. Buller is explained afterwards. Finally the modeling of the defense mechanisms and the implementation of them is explained. In the next chapter an overview of the state of the art is given.

2 State of the Art

Various bionic approaches of describing the human brain and behavior are possible. One feasibility is a neuroscientific model, which provides concepts for data perception and symbolization, and serves as a template in form of a bottom up design. As well, in psychology many varying, partly conflictive models exist, which are describing the human behavior from different points of view. The psychoanalytical model of the human mind is one of those models, which was first initiated by the basic research of Sigmund Freud and has been essentially improved within the last 20 years. In the project ARS (Artificial Recognition System) a unified model of these two different bionic approaches has been used [BLPV07]. In the following, different emotion-based approaches to artificial intelligence and the model of A. Buller are explained in detail. A. Buller uses psychodynamics to describe the human mind and implements his concept in Volitron, a controller that should increase the competence in such activities as self-initiated exploration of an environment. In his approach he also applies first concepts of some defense mechanisms. Afterwards the Bubble Family Game (BFG) is described, which is the first prototype of the first model of the project ARS explained in detail in Section 3.2.1 and is used to implement the defense mechanisms. Finally the key results of the diploma thesis of S. Kohlhauser [Koh08] are presented, who had done a requirement analysis for describing use cases for testing the behavior of the agents simulated in the BFG.

2.1 The Theory of Emotion-Based Approach to Artificial Intelligence

The opinion that emotions are an important facet of intelligent behavior was accepted recently in the research of Artificial Intelligence (AI). The neuroscientist Damasio is convinced that emotions are crucially intertwined with cognitive problem solving and decision making. Also the psychologist N. H. Frijda argues that emotions are an essential part for the establishment of social behavior. Because of the main problem to reach a profound comprehension of emotional behavior, the task of defining emotions in technical terms is seen as difficult and has often been accounted as not feasible. Nevertheless, an increasing number of researchers believe that it is needed to develop a computational model of emotions for designing an intelligent autonomous agent in general as well as to create a new generation of robots which are able to socially interact with each other or people in special [PB09].

M. Minsky wrote in [Min86, p. 163]:
"The question is not whether intelligent machines can have emotions, but whether machines can be intelligent without any emotions."

During the last years, there have been developed various kinds of systems trying to implement and investigate emotions. They differ significantly concerning to their aims and assumption as well as refer in several degrees to existing theories of emotions. The concept of emotion is very broad and various aspects including physiological, motivational and expressive ones are covered. Also the subjective experience of emotional states in the form of feelings and the ability to cognitively reason about emotions are involved. So far, only a subset of these aspects are included in the most of computational emotion systems. The existing work can roughly be divided in 'communication-driven' and 'process-driven' approaches. The first approach is focused on the surface manifestation of emotion and their influence on human-computer interaction, the second one tries to model and simulate the mechanisms of emotions as they unfold. But there also exists some systems which address both perspectives. In the following, a small number of key projects mainly focusing on the process perspective of emotions will be described [PB09]. Also the model of A. Buller, who uses a psychodynamic concept is presented.

2.1.1 Ethology-Inspired Models

It is clear that emotions possess bodily aspects. To protect body's integrity and well-being, the evolution of emotional mechanisms has been a primary task. Emotions, which have been seen as control mechanisms, that have been governing behavior long before higher-level cognitive mechanisms showed up, are often utilized when constructing selection architecture for autonomous adaptive agents. This approach leads to ethology-inspired architecture. For this architecture agents with a physically embedded body - whether real or simulated - and mechanisms to prioritize the use of limited resources are important design aspects. An example therefor is the model of Canamero explained in detail in [Can97], in which robots have the task to survive as long as possible. The environment, in which the robots live, contains various types of resources, obstacles and predators. To maintain the well-being, some different actions are to carry out by the robots. The behavior selection architecture of the robots includes the components *physiology*, a set of *motivations*, a set of *behaviors* and a set of *basic emotions*. *Physiology* consist of synthetic survival-related variables like 'energy' or 'blood sugar'. *Motivations* are for example hunger and aggression, *behavior* are for instance attack and eat, and samples for *basic emotions* are fear, anger and happiness. If a survival-related variable depart from their homeostatic regime, motivations are activated e.g. when a robot is too warm, its motivation to decrease its temperature is invoked. Each motivation has an intensity which is influenced by the emotions of the robot and the motivation with the highest intensity controls the behavior. Trigger for emotions can be the presence of external objects, or the occurrence of internal changes or pattern. An example for an internal pattern is when one or more of the motivations is too high, wherefore the robot becomes angry. Activated emotions initiate hormones that have an effect on the physiology, attention and perception of the robot. According to Canamero, this model incorporates only simple and low-level mechanisms and that issues such as learning is not addressed. However, different reward and punishment mechanisms can be accommodated into this model. Both are very important for learning which argues Canamero as well as other researcher.

2.1.2 Emotion-Related Learning in Autonomous Agents

Because of showing organism what is 'good' and 'bad' for itself, emotions can be used for learning. Learning, which is based on positive or negative feedback, is named reinforcement learning. It is widely used, however without referring to the concept of emotions. But the implementation of emotional mechanisms is argued to provide additional flexibility. O. H. Mowrer demonstrated that emotions enable a two stage learning process. The first step is a response to a special stimulus with a special emotion, which is learned by the agent. The second step is a behavior learned by the agent, which is associated with an influence on the emotional state. At the same time the implementation of emotional states animates the agent to seek for different meanings of behavior, if the old one is not useful anymore. Furthermore, explicit emotional states can influence more than one process simultaneous. An example therefor is the connectionist model of emotion synthesis and learning from J. Velásquez, which is called Cathexis and described in detail in [Vel98]. He physically implemented it in a robot dog called Yuppy, which has six basic emotions and contains a perceptual, a drive, a behavior and a motor system as seen in Figure 2.1. There are four types of releasers of emotion, e.g. neural elicitors like neurotransmitter or motivational elicitors like drives. The calculation of the emotion happens separately, but the update-rule has the same form for each type of emotion. The new intensity of each emotion is a function of its elicitors, its decayed previous value and influences of other emotions. It is compared to an emotion-specific activation threshold and only when this threshold is passed, the emotion influences the behavior system as well as other emotions. Emotions can be a result from interactions with the drive-system, the environment or people. Yuppy is able to learn in the form of 'secondary' emotions, when rewarded or disciplined by persons. For example, Yuppy can learn to fear the sound of a flute when it occurs often enough together with pain. This model remains rather still low-level because of focusing on basic control circuits implemented via connectionist networks and symbolic representations are not explicit used.

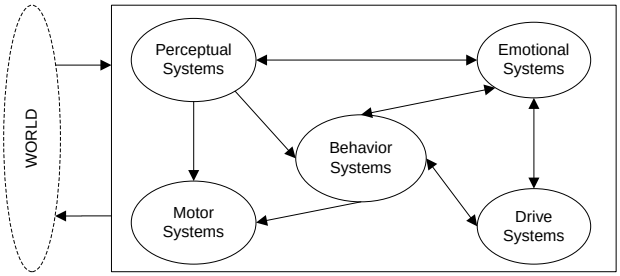


Figure 2.1: Schematic of the Framework for Emotion-Based Control [Vel98]

2.1.3 Appraisal-Based Models

In the appraisal-based model the focus is solely on the cognitive structures and mechanisms, that are included in the generation of emotions. Affective reactions are the result of cognitive appraisal processes which map the features of a situation onto a set of output emotions. Therefor high-level rule-based representations of goals, preferences and situations are utilized. The OCC model, developed by the emotion theorists Ortony, Clore and Collins, is one of the early appraisal-based model and has served as the basis for implementation of several models of emotion. It has not been designed for synthesizing emotions, but to enable AI systems to reason about emotions,

which is seen as a useful capability for applications such as natural language understanding and dialog systems. In this model basic emotions are not used, but emotions are grouped according to a scheme of cognitive eliciting conditions. Three classes of emotions are distinguished, which depending on the stimuli that causes the emotions - those induced by events, those induced by agents and those induced by objects. The feature of situations are linked with the belief and the goal of the agent by a complex set of rules. For that intervening structures and variables are used, e.g. if $D(a, e, t)$ is the desirability that an agent a attachs to an event e at the time t then the capability for generating a state of joy P_j is given by a joy-specific function. This function depends on the assigned desirability of the event and a combination of some global intensity variables, e.g. expectedness, and is represented by $I_g(a, e, t)$.

$$IF \quad D(a,e,t) \geq 0 \quad THEN \quad set \quad P_j(a,e,t) = f(D(a,e,t), I_g(a,e,t)) \tag{2.1}$$

The rule seen above does not directly activates a state of joy, but another rule, which causes joy with a certain intensity I_j only if a joy-specific threshold $T_j(a, t)$ is exceeded. In the model details like what values to use for threshold or how emotions interact, mix, or change their intensity, is not very specific and it is not the most flexible one. It is basically a knowledge-based system to generate different types of emotions. Almost no feedback from the emotional system to the cognitive system is implemented [PB09].

2.1.4 Architecture-Level Models

Architecture-Level Models try to combine low-level mechanisms of emotion with schemes of how to cognitively elicit emotions. A prominent prototype of such a model is the approach of A. Sloman and his colleagues to emotions and intelligence. However, he is not particularly interested in emotions, at least not in creating artifacts that just simulate the effects of emotional behavior, but in construction of a general intelligent system. Moreover, such a global architecture of the mind also has to contain mechanisms for perception, learning, making plans, drawing inference, etc. All these mechanisms have to be implemented in an agent.

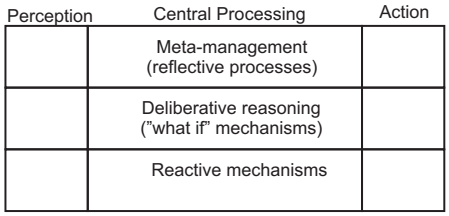


Figure 2.2: The CogAff Schema [SCS05]

During the last year, various drafts for a cognitive architecture have been proposed by them, but all of them are based on an architectural schema that is very generic. It is called CogAff scheme and is able to integrate various kind of emotional and non-emotional mechanisms. Sloman believes that a cognitive architecture for the human mind needs at least three layers, which he named reactive layer, deliberative layer and self-monitoring layer as seen in Figure 2.2. The reactive layer can release only relatively simple and predictable behavior and is based on the detection of characteristic stimuli in the environment. Rather automatic behavior responses are produced, e.g. reflexes based on direct connection from sensor to motor are the most basic form of reactive

mechanisms. The deliberative layer involves 'what-if' mechanisms and contains formalisms to combine existing behaviors in new ways, making plans, describing alternatives and evaluates them before executing. The meta-management layer allocates mechanisms for monitoring and evaluating internal states and processes. In order to keep processes of the lower layers from interfering with each other as well as to increase their efficiency, the meta-management layer is able to control, reject, modify and generalize them [SCS05].

In the Sections 2.1.1 to 2.1.4 various approaches for modeling emotions are illustrated. As seen, emotions are important for our well-being and help us to learn what is 'good' and what is 'bad' for our organism. In the next Subsection the model of A. Buller is explained, who uses psychodynamics. In his paper [Bul05] he describes initial ideas and interprets the psychodynamic concept in more technical terms, to make them useful for building artificial agents. He also implements his ideas in robots, like Miao. In Subsection 2.1.5 a short overview about his ideas and the implementation of Volitron, a controller, which should increase the competence in such activities as self-initiated exploration of an environment, is given.

2.1.5 An Approach using Psychodynamics

In his paper [Bul05] A. Buller uses concepts of psychodynamics for his model and argues that Artificial Intelligence (AI) could benefit by dealing with some Freudian concepts. Important for Buller is that he uses psychodynamics and not psychoanalysis. For him psychoanalysis is a theory and therapeutic method based on the following four assumptions, but psychodynamics only accepts the first two of it:

- The fundamental role of unconscious processes
- The existence of conflicting mental forces and defense mechanisms
- The existence of the Oedipus complex
- The key role of sexual drive and aggressive drive in the development of personality

But the definition of the psychodynamic approach is not universally. For example M. Jarvis proposes that primacy of affects, the continuity between childhood and adult experience and so on should be assumed by psychodynamic approaches.

To be able to use psychodynamic concepts, he interpreted these concepts in more technical terms. Some of his initial ideas therefor are explained in the following:

- *Tension* is a physical quantity. It is associated with a physical or simulated tension-accumulating device, in which the tension is accumulated or discharged as a reaction to an input signal.
- *Thoughts* are meaningful patterns, which are presented by the states of a tension-accumulating devices.
- *Feelings* are subjective experiences, which an agent has when e.g. elements of its working memory get important bodily signals.
- *Pleasure* is one of the feeling, which emerges during a certain tension is discharged.
- *Wishes* are meaningful thought-like patterns, which are saved in the working memory. They are representing their imaginary object or situations, which possible causes pleasure.

- *Conflicts* occurs if two or more contradictory wishes appear, whereas the agent can only do one action in favor to only one of them.

He also tests them in simple robots. Simple two-motor vehicle are used with camera and speaker, to demonstrate how tensions and pleasure work. His first experiments are the robots Neko, Miao and Miao-V, which use the tension-driven behavior. Furthermore he designed a controller named Volitron, which is explained in detail in [Bul02]. This controller should increase the competence in such activities as self-initiated exploration of an environment, new goal acquisition, etc. of its host robot. The structure of Volitron shown in Figure 2.3 includes following four key elements:

- Model of perceived reality
- Model of desired reality
- Model of ideal reality
- Model of anticipated reality

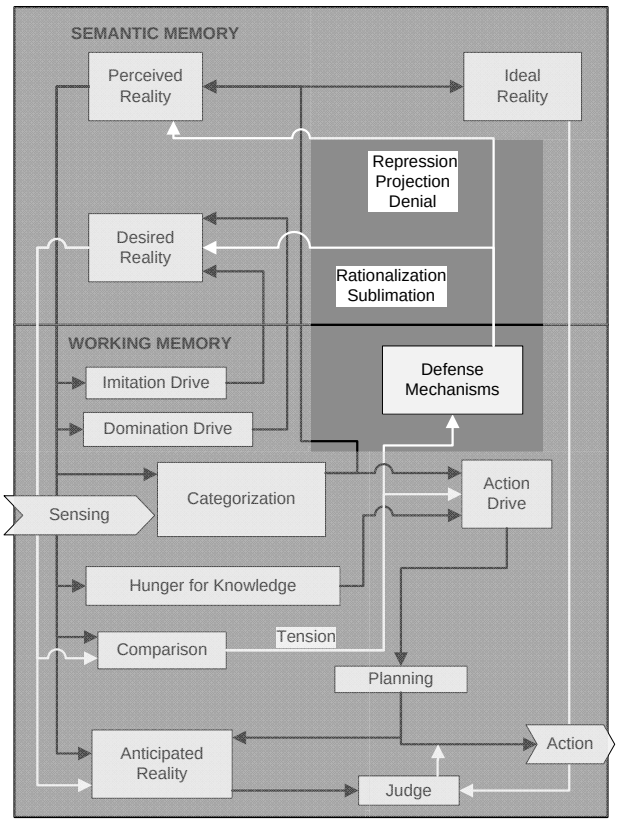


Figure 2.3: Structure of Volitron [Bul02]

As shown in Figure 2.3, he also applies the five defense mechanisms Repression, Projection, Denial, Rationalization and Sublimation. These mechanisms starts working, if one of the following two situations has happened. The first situation is that no satisfactory plan has been generated for a long time. The second situation is that the already executed plan could not change the environment so as to reduce the tension. Pieces of information are called memes and can be produced by defense mechanisms. They are able to cause changes in the models of reality. An

example for the model of perceived reality are the changes, which manifest themselves as a repression of inconvenient facts. Due to these changes, the tension will eventually be reduced in some way.

However he also compares this model with psychoanalytical terms of the second topical model of Sigmund Freud (detailed explained in Section 3.1.1). In the model of Volitron, the perceived reality together with its devices for planning and judgement can be seen as a counterpart to the conscious part of the Ego and the defense mechanisms may be understood as a unconscious part of the Ego. The model of desired reality along with the devices for drives and tension creation can be interpreted as a counterpart to the Id and the model of ideal reality as an artificial Superego.

2.2 Bubble Family Game

The first prototype of the ARS-PA (Artificial Recognition System - PsychoAnalytic) architecture, which is explained in Section 3.2.1, is called Bubble Family Game (BFG) [DZL07]. It is written in Java embedded in the simulation tool Anylogic[1]. The BFG is an object of study and not supposed to fulfill a particular need like serving as an alarming system. It is a simulation of artificial life and implements following three elements:

1. the body of the artificial creature,
2. the environment, and
3. the mind of the artificial creature.

The *body* of the artificial creature is a rather simplified one. The artificial creatures are called *bubbles*, whose task is to survive as long as possible. In spite of the simplicity - physical dimensions and structures are not considered in this first prototype, the body is extremely important for the occurring drives and emotions, for the grounding of the symbols, and for developing intelligent behavior in general. In contrast to that R. Pfeifer wrote in [PB07] that in our embodiment lies the foundation of the kind of thoughts we are capable of, which means in our morphology and the material properties of our bodies. This was also seen in the project ARS-PA and therefore the physical dimensions and structures of the agents are reconsidered in the new simulator. The body features of the bubbles are:

- Proximity sensor and eyes
- Movements
- Eat from energy sources and
- 'On-board' energy storage

The process of symbolizing is the transformation of the sensor values into meaningful chunks and is described in [Pra06] in detail. Symbols which a bubble gets are for example <energy source is in front> or <other bubble on the left> and the distance of them. A bubble also has a homeostatic level. A kind of 'battery' must be kept within a certain level and actions like walking drains it. To refill it, energy sources have to be used. Bubbles are able to act out four modes of operation, which are named promenade, attack, flee and eat.

[1] http://www.xjtek.com

Also important is to put the artificial creatures into a setting that stimulates their senses and that allows for interaction with the surrounding and with other bubbles. The BFG has an *environment* which is a 2-dimensional, finite playground with a number of different energy sources and number of other bubbles. This setting might look trivial, but it is necessary to provide a stimulation of a certain analyzable and feasible complexity to the bubbles.

The actually most interesting part of the simulation is the cognitive system of the bubble - its *mind* - which implements:

- Drives: hunger, seek and play
- Basic emotions: fear, rage and lust
- Complex emotions: hope, joy, disappointment, gratitude, reproach, pride, shame, etc.

The realization of the level of drives and emotions are lean on the real world, in which they have various characteristics with a ranking relationship, but cannot be said in numbers. An example therefor is hunger, because a human is only able to say that he/she has low or high hunger, but the difference between low and high cannot be defined respectively cannot be calculated. To be able to use these characteristics for computing, every characteristic is enumerate by numbers, whereas very low, the minimum, is number one and very high, the maximum, is number seven. Thus, it is possible to compare, if an emotion or drive has exceeded a threshold.

There are two different ways, how drives can be influenced by the properties of the body. The first and most used one is the directly connected, e.g. the hunger because if the energy level drops below a certain threshold hunger rises. The second one is that drives rise when a mixture of internal and external conditions applies. So far, there are two categories of desire. The first one only deals with individual bubbles and objects, e.g. a bubble can get desire for a very specific energy source. The second one, being most of desires, deals with bubbles as being part of social group and are supported by some rules. For instance some energy sources are only consumable when more bubbles working together or some bubbles desire a group to play. The purpose of the rules is, to evoke situations which require cooperative behavior. So the interest is on the 'configuration' of the bubble such that they can solve problems or task as a team. For bubbles, social acceptance and achieved tasks are pleasurable states. For solving cooperative tasks, previously experienced scenarios like 'I am hungry but the energy source is too big to be hunted down alone' have to be remembered. Main building blocks of the action planning procedures are:

- Images, which are snapshots of a set of recognized symbols
- Scenarios, which are sequences of images
- Desires, which are activated wishes to re-experience once pleasurable scenarios. The emergence of desires is explained in detail in [LZD+08]. The satisfaction of them discharges a tension in a certain amount

A recognized scenario can activate an associated desire and sequences of scenarios serve as a blueprint for a bubble how to satisfy directly or in several steps the desire's tension. The desire is stored with a 'chance' of success which influences the pursuit of the desire and this change is set to 100%, if the desire is finally satisfied, and the tension is discharged. Currently scenarios and desires are stored within an episodic type of memory, and if they are not needed for a certain time, a bubble can gradually 'forget' them, which means that the significance of it reduces with time.

Scenarios and desires are more likely remembered, if their emotional impact on the bubble is big. Each bubble possess its own, personal *Episodic Memory* module in which all its experiences and previous adventures are stored. The definition of the *Episodic Memory* is done as directed graph in XML-format. The *Episodic Memory* module has two parts, both of them consist of images as states but they differ in the edge. The edges of the Scenarios are *images* that match but the edges of Desires can either be *images* or finished *actions* that match. When a certain image - the init STATE - is recognized, both graphs can be 'activate'. After all states are traversed, the sequence is fully recognized and the graph is deactivated. However, a scenario graph can not only activate a desire, but can also increase the tension of an existing, active one. As well it is not unusual that several, eventual similar, scenarios are simultaneously activate because the activation is done by the exceeding of a certain confidence level which indicates the similarity of a stored template image with the currently perceived input. Complex Emotions can be caused by the states of an active scenario or an active desire graph. Both types of graphs have timeouts. Either this causes a deactivation of the particular graphs, if they stay too long in one state and the timed-out desire/scenario graph leads to frustration or it falls back to the last state after time out. The implementation of the Episodic Memory is done as an associative memory, which means that the entry X can be searched by its meaning and not by its address. The search of an entry in the Episodic Memory is flexible and fuzzy. For example if it is searched for 'a cup falling from a table' scenario, several different scenarios, like one where it was fell down and broken, where it did not break and where it was caught by someone, are retrieved. Currently there is no explicit sense of time within the model, but it exist the timely order of nodes. This kind of implementation of the Episodic Memory enables the bubble to select its action to take [PPC09].

One concurrent block of functionality is being implemented in the entire simulation loop and consists of eight parts [PPC09]:

1. **Perception:**
   The composition of symbols to images is done, according to [LBP+07]. There is no work on the level of sensor data.
2. **Image matching:**
   After comparison of the sensed image with stored ones (template images), which are the knowledge base of the agents, an ordered list from 0% to 100% is formed. The respective recognition level is increased by further 5% if the system is waiting for a certain image, which can be seen as a simplified implementation of 'a focus of attention'. A reactive action tendency can eventually be issued by all matching images, which are sorted by matching level, and importance (e.g. recognition level above 70%). Drives and basic emotions are directly activated by images, which can even cause contradicting ones.
3. **Scenario Processing:**
   If the condition for a state change (=edge) of an active scenario graph is fulfilled, the state is changed and if a special 'INIT image' is recognized, the corresponding INIT state is invoked, i.e. the scenario is activated. Completely recognized scenarios can activate a desire.
4. **Desire Processing:**
   Desires are similar to scenarios. The difference to them is that now open nodes can create desire action tendencies. They are sorted by their current intensity of the associated tension, and their likelihood of success. As well an update of the level of the complex emotions is done.
5. **Merge of action tendencies:**
   The merge of the sorted lists of all existing reactive and desire action tendencies is done.

6. **Superego:**
   The Superego takes influence on the action tendencies and simply does not allow some of them. As well new action tendencies can be generated. At the moment it is very rudimentary, because it is only implemented that if one scenario is active another one is not allowed, but it will later deducts social rules out of the current scenario.
7. **Conflict detection and resolution:**
   The finalization of the ordering of action tendencies from all sources is done, whereas different sources can be differently weighted. If a conflict is still there after this - it could be possible because of the fuzzy/integer representation of action tendencies - a decision is made by random selection.
8. **Execution:**
   Currently, only one action tendency per simulation step and bubble can be executed in the BFG.

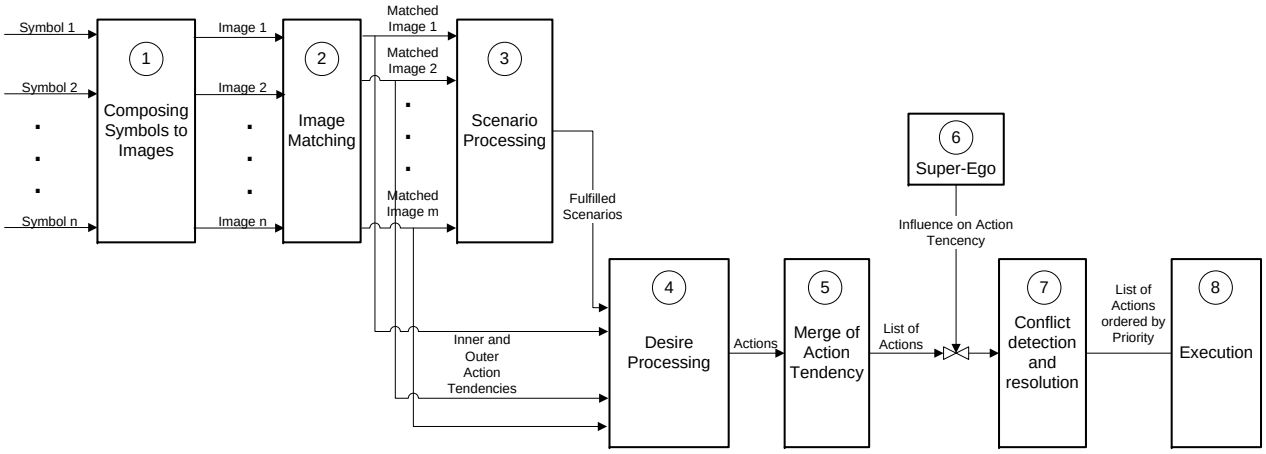


Figure 2.4: Block Diagram of Functionality within the Implementation

The Architecture of the Simulator explained in detail in [RLD+07] can be seen in the Figure 2.5 and consists of the following three modules:

- **World Simulator:**
  This module includes the environment and the basic knowledge about every agent and has following tasks:
  - Collecting action commands from agents
  - Verfing these commands due to simulation rules
  - Executing the actions
  - Collecting sensor data for each agent
  - Transmitting these data back to the agent
- **Agent Body:**
  This module operates as an abstraction layer for decoupling the interface to the *Behavior architecture* from the interface to the *World simulator*. It can be seen as the interface between the *World Simulator* and the *Behavior Architecture*.

- **Behavior Architecture:**
  This module contains the decision unit of the agent and defines its behavior, in the BFG e.g. attack or flee. It is put into a separate module for each agent, to enable the usage of various types of realization of behavior.

Due to this architecture, it is possible to simulate different types of agents with different behavior architecture in different worlds as well as to distribute the simulation on several computers. The simulation platform is explained in detail in [DZLZ08].

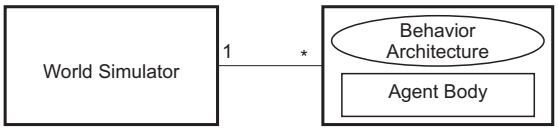


Figure 2.5: Simulator Architecture [RLD+07]

2.3 A Psychoanalytically Inspired Agent Based Social System

An approach for designing the decision unit for autonomous agents is using findings of psychoanalysis. The team of the project ARS develops a model of the human mind explained in detail in 3.2.2, which is based on the second topographical model of S. Freud. This model should be implemented in an Artificial Life system like the BFG, which is the first archetype of the first model of the project ARS. To be able to observe the behavior of the agents, Kohlhauser has described some behaviors and defines some use cases for testing these behaviors in [Koh08]. The description of the behavior is based on findings in general psychology and social psychology. In the following a short overview of different behaviors is given. This description is the theoretical background for the use cases explained in the next chapter and is done by means of their real world occurrences ([Koh08, p. 29-40]). These behaviors and its changes are analyzed in the next chapter, because it differs due to the usage of defense mechanisms. This Section describes how the defense mechanisms influence the behavior of humans. The changes of the behavior are illustrated by use cases, which are explained in detail in Section 3.3.3. One point, which is important for observing behavior is, that it can not be predicted deterministically. Every single decision is influenced by various factors. Some decision are rational and can therefore be understood easily. As soon as complex reasoning based upon affects, emotions and feelings is involved, a prediction of the taken actions becomes difficult [Koh08, p. 41].

Love

It can be distinguished between two kinds of love - the *passionate love* and the *companionate love*. The basis of the first one is the physiological excitement and has as result the fulfillment and ecstasy if it is return. If it is not returned it may cause sadness and desperation. Various factors have an influence on the feeling of attraction towards another person, which raise the possibility to be good friends or to fall in love. One of them is the *effect of closeness*, which means that someone tends to like persons, who he/she often sees and interacts with. Another factor is the similarity e.g. in appearance, beliefs and values. This is important because of increasing the change that the feeling is returned by the other person and affirms our beliefs. Other factors are

reciprocial affects and *bodily attractiveness*, whereas the first one means that if someone likes a person, he/she tends to return the affection.

In the case of not returning the love by the other person it can be better to repress or sublimate this drive, which is done by one of the defense mechanisms. The reason therefor is that sadness or desperation may lead to aggressive behavior, which may not be accepted in the group and can have as result the excluding of the group. The use of defense mechanisms depends on the intensity of the feeling love and of the self-esteem of the person. If the feeling is not strong enough it may have no effect on the person because of searching for another one which may return the feeling.

Aggression

It can be distinguished between three kinds of aggression - the *hot aggression*, the *cold aggression* and the *male dominance behavior*. But in the following all descriptions refer only the *hot aggression*, which is defined as "intended behavior, with the goal to afflict other people with physical or psychic harm" by Aronson [AWA04, p. 440]. Aggressive behavior can be caused by various factors e.g. *pain*. As well environmental conditions, which let someone feel not good like high temperature or loud noise, raise the possibility of aggression. Another factor is *frustration*, which emerges if someone are prevented to reach a goal important for him/her. The intensity of aggression caused by frustration is dependent on the possibility of reaching the goal as well as whom is disturbing him/her to reach the goal. Because if the opponent, who prevent someone reaching the goal, is stronger he/she may not become aggressive against him/her.

Sometimes it may not be advantageous to show aggressive behavior in the own group. Defense mechanisms can help to repress or sublimate these drives or make it possible to delay the discharging or displace the person, on whom someone is angry, by an object which has nothing to do with the primary situation. As well if the opponent is stronger as oneself it also may be better not to be aggressive at that moment and so it should be possible to discharge the aggressive drive at someone or something else, which also can be done by a defense mechanism e.g. Displacement.

Stress

Stress is caused by a stimulus, which can have its reason named stressor internal as well as external. This stimulus disturbs the homeostasis of the organism and strains or exceeds the ability of the organism to deal with these effects. It can be distinguished between two kinds of stress - *acute stress* and *chronic stress*, whereas the first one is only temporary and the second one is a continuous excitation. Chronic stress shows the individual that the internal and external resources do not fit for dealing with the situation.

If a person has chronic stress, he/she goes through three stages. In each stage it is possible to use different defense mechanisms. The first stage is a short phase and is called *alarm reaction*. In this phase the body prepares for energetic actions. In this stage defense mechanisms are used for repressing drives which are not needed at that time because of needing all the resources of the body for dealing with the stress. If the stressor continues, the body enters the second phase named *resistance*, in which the body is able to defy the weakening effect of the stressor. If the stressor persists for too long or is too intensive, the resources of the body are exhausted and the body enters the third phase named *fatigue*. A permanent state of stress can make the immune system weak and the body susceptible for disease. In this phase the body may use the

defense mechanism Conversion because the homeostasis of the organism is too much disturbed and searches for a possibility to show the person that the body is too weak to do anything and that no drive discharging is possible to balance the needs of the body.

Various categories of stressors are described in the literature. One of them is *big life changing events*, which can be positive e.g. a baby or negative e.g. losing the job. Another one is *catastrophes and traumatic events*. In this case the body uses defense mechanisms during the day to minimize the stress during the day, but in the night the traumatic event is undergone in dreams. The most common cause of stress are *everyday problems*, which are for example being late, loud environments or forgetting something.

Reciprocity

In the theory of reciprocity it is argued that someone feels obliged to pay back a favor to someone else, who does him/her a favor, whereas this happens independent of the sympathy between both. If he/she does not return the favor or gift it may be a disadvantage for him/her because the other will despise or avoid him/her and it is less likely that the other help him/her in the future.

It may be possible that someone does not trust the other. Nevertheless he/she returns the favor because of believing the other one may help in dangerous situation although knowing that this would not happen. To do so defense mechanisms are applied, e.g. repressing that the other would not help, or giving a logical reason like that the other will help him/her if he/she is in affliction, which is done by using Intellectualization. A reason to act so can be that someone is falling in love with the other.

Hierarchy

The simplest hierarchy has two levels, whereas the leader is at the top and the group members are on the lower level. The leader is usually responsible for making most of the decision for the group. In the following all descriptions refer to the pecking order within a group.

In a hierarchy the authority and the influence of a member within the group is defined. In the group some members have power over the means of need satisfaction of other members. These individuals have leverage by different resources (e.g. food) over these members, because if these members have e.g. a large quantity of goods they can reward or punish the others.

In a hierarchy some members have no or only a small influence on the decision, which goals have to be reached. This can have as result that their drives become not satisfied if they differ from the drives of the group. Therefor it can be necessary that the agent has the possibility to repress its drives to be able to follow the goals of the group. A reason for following the leader may be the shortage of a resource, which the agent may receive as reward from the leader and helps the agent to survive.

Sometimes a leader pursues his/her goals in a hard way, which means that he punishes the members of the group very hard if they do not follow his/her commands. If the reward is not high enough it may be difficult to motivate the group members to help the leader for fulfilling his/her plans. But if a person is falling in love with the leader the defense mechanisms may influence this, e.g. the repressing that the behavior of the leader is not fair which is done by Repression, or giving himself/herself the reason that it is necessary to follow because of surviving which is done by Rationalization.

Ability to Work in a Team

To be able to build a team, it is necessary to have the ability to work together. This can be interesting because some tasks are better done in a team, e.g. to fight against another team, which attacks the agent. Sometimes it can happen that an agent only sees another agent, which it does not like, but needing help for an important task or is asked for help. Therefor defense mechanisms can help to ask someone for help or to help someone although not liking him/her. This can be necessary e.g. for surviving or to make someone a favor to become help when needed.

In this chapter a short overview about different approaches to use emotions in autonomous agents is given. The model of A. Buller, who uses psychodynamics in his concept, in which he also implements four defense mechanisms, is described. Afterwards the Bubble Family Game (BFG) is explained, which is the first prototype of the first model of the project ARS, and is used to implement the defense mechanisms. Finally human behavior according to [Koh08] and reasons for using defense mechanisms is described. In the next chapter the psychoanalytical theory about defense mechanisms is presented, to analyze the various mechanisms and find out some important characteristics. Afterwards the two models of the project ARS are explained. By a view on human usage of these mechanisms a general model is developed, which is transformed into a technical model for realizing them in the BFG. Afterwards this model is implemented in the two models of the project ARS and the information flow is presented. Finally a general model of the four defense mechanisms Displacement, Sublimation, Intellectualization and Conversion is described.

3 Theory and Model

The main part of this work is to define a technical model for psychoanalytical defense mechanisms, which are important in the human mind to deal with an inner conflict. An inner conflict means that two contrary drives or desires occur and both would like to become satisfied. Because it is not possible to fulfill both, human would become frustrated, which is caused by a conflict between the instances *Ego*, *Id* and *Superego* - the three instances of the second topographical model of S. Freud. To avoid such situations defense mechanisms become active. In this chapter the psychoanalytical theory like the Structural Model of S. Freud and various defense mechanisms are explained first. As well the motivations for using defense mechanisms in the human mind are presented and why they are important in autonomous agents. Afterwards a general human model for the defense mechanisms is described, which is transformed into a technical model to be able to implement them into the Bubble Family Game (BFG). Finally, some defense mechanisms are analyzed in detail, e.g. important characteristics for technical realization are explained.

3.1 Psychoanalytical Theory

In this Section the Structural Model of S. Freud (Ego, Id and Superego) is described, because the theory of defense mechanisms is based upon it. This model is called second topographical model. In the first topographical model he distinguished between the three agencies Unconscious, Preconscious and Conscious. Freud attended to the theory of drive defense in [Fre40c], which was the base of research for the defense mechanisms. He differed between Reversal, Turning Against the Self, Repression and Sublimation, whereas he concentrated his research on the mechanism Repression. This term was used instead of defense for a span of time. But later it was seen that Repression is one possible defense mechanism and for the general description the term Repression has been replaced by the term defense. His daughter A. Freud analyzed various defense mechanisms based on the research of his father in detail. These mechanisms are explained afterwards, whereas the analysis of them is done by means of human examples to obtain an overview of lots of different defense mechanisms and to be able to create a technical model for implementation. Furthermore the reasons, why a human utilizes defense mechanisms, are illustrated and are used for derivation of reasons, why they could be meaningful in autonomous agents.

3.1.1 The Second Topographical Model

The Structural Model [Fre40a] seen in Figure 3.1 is S. Freuds second topographical model of the mind, in which he highlighted three various psychic entities. They are called *Id*, *Ego* and *Superego* and are in constant contact with each other. The *Ego* is those instance which has contact with stimuli coming from outside. It is the core of the model, which has a leading role and is responsible for the homeostatic regulation with regard to the instinctive drives of the *Id*. The *Ego* deals with the *Superego* which is a structural and mainly unconscious part of the individual's personality. The *Superego* develops in early childhood and contains moral rules, which often serve judge, inhibit and punish but which are occasionally also a benevolent guide and the source of narcissistic approval. The Superego is just like our inner voice, which tells us: "You're like this... but you ought to be like that..."[PPC09, p. 8]. If the compared alternatives are too distant from one another, frustration and a conflict within the person, who develops an anxiety signal of varying intensity, are the result. The hub of Freudians Structural Model is the *Ego*, which has two functions. On the one hand, the *Ego* is the executive agent, urging, planning and organizing voluntary actions. On the other, the *Ego* is an agent controlling and inhibiting the senses. Both is done with respect to the outer world and the inner world particularly with regards to wishes and instinctive and motivational urges. The *Id* is a part of the unconscious system and can be viewed as a boiling container of experiences, thoughts and fantasies that are repressed because they are unbearable to the conscious mind. However, *Id's* repressed experience locked in area that are not subject to the law of logic. For example in the *Id* there is no room for denial or for mutual contradiction, in which two contradicting phenomena can coexist. There is also no law of time in which passage of time is entirely unacknowledged. According to Freud, the atemporal organized *Id* explains the enduring influence of the repressed experience on observed subjective behavior [PPC09].

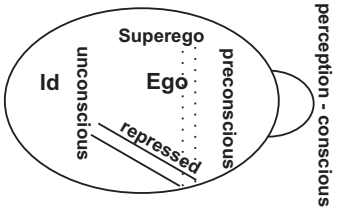


Figure 3.1: S. Freud's Second Topographical Model [Fre33]

3.1.2 Psychoanalytical Defense Mechanisms

The term of defense is defined as the resisting of the Ego against embarrassing and forbidding conceivabilities and affects. At first it was used in the 1894 in the study of 'Die Abwehr-Neuropsychosen' [Fre40b, p. 57-74] and has been since then the oldest representated dynamic concept in the theory of psychoanalysis. Later the term repression was used but in the 1926 defense was taken on again and repression became a specific defense mechanism. So the new challenge was to find more other defense mechanisms. One reason therefor was the discovered context between some affects to the defense mechanisms; for example Regression, Reaction Forming, Isolation and Undoing often be used by people who suffer from compulsion neurosis. In some other works of S. Freud defense mechanisms like Projection, Introjection and Identification, Turning Against Onself or Reversal have been described. All of them are functions of the Ego, because

without an appeal of the Ego or of the outer world, which allow the Ego to act for, drives would only experience one destiny, namely satisfaction. The last defense mechanism which was mentioned by A. Freud is Sublimation, which causes a displacement of the driving goal. According to A. Freud, the Ego possess these ten defense mechanisms to struggle against drive representation and affects at disposal [Fre84].

In [Fen97] O. Fenichel tries to devide the defense mechanisms in successful defense, which knocks the defensed on the head, on the one hand and the unsuccessful defense, which makes the repetition of the transaction necessary, to inhibit the break out of the drive impulse on the other hand. But the boundary between both is not strictly, because occasionally it is impossible to recognize a difference between an impulse, which is changed by influence of the Ego, and an impulse, which breaks through against the will of the Ego because of distorted form being unrealized. The latter one normally evokes a cramped behavior, which repeats ever and ever again. Also a complete relaxation is not allowed and causes a symptom of fatigue.

In [SSK97] the arrangement of defense mechanisms by Paulina Kernberg is used. She differs between normal, neurotic, borderline and psychotic defense mechanisms. The defense mechanisms described by A. Freud are assigned by her to normal and neurotic ones, but she also describes further ones.

The following defense mechanisms are described in [Fre84] or in [SSK97]. The explanation of the defense mechanisms of O. Fenichel in [Fen97] is not introduced, because he has a different categorization of them.

Quotation of defense mechanisms:

Repression
While I repress something, I do not look at it, I do not see something and I do not want to think about it/ have it in awareness [SSK97].

Repression is an action of the Ego, with which particular drive wishes connected with mental content (associations, feelings, ideas, phantasies etc.) are repulsed or are held off from awareness. The consequence of repression is that in his/her conscious experience/life nothing of the content is detected and in his/her subjective experience he/she thinks that incidence never happened.

Denial
I am introverted [SSK97].

Denial is a defense mechanism by which a person tries to eliminate an awkward or undesirable situation of the outer reality within its traumatic meaning with help of a fulfilling phantasy or by an outer behavior. For example after the school an eight year old boy sees on the way at home how his best friend dies. He denies this situation and would like to meet his friend next day in the morning to go together to school.

Displacement
I redirect my feelings and emotions on objects which are not concerned with them.

If Displacement is used as a defense mechanism, people, who have unacceptable feelings, emotions, wishes or drives directed on one object, redirect them onto another object. For instance, someone punches a cushion when he/she is angry at his/her friends. Another example is when someone has problems at work and cannot speak with his/her responsible boss about it, he/she takes his/her anger at home and argues with his/her partner.

Sublimation
I stand above the conflict, while I cope with it [SSK97].

Sublimation is used when unacceptable impulses e.g. aggressive drives appear. Therefore this behavior is converted into a more acceptable form. For example someone takes his/her aggressive impulses out on sport like boxing or intellectual activities.

Projection
I split, leave out, eliminate [SSK97].

By Projection someone ascribes his/her feelings and emotions to other people. For example someone does not like a person but thinks that the other person does not like him/her.

Rationalization
I offer possible reasons instead of the real one [SSK97].

When Rationalization is used as a defense mechanisms, people offer not the truth reason why they do something. For instance someone is in love and does know that the loved person needs a book. So he/she brings him/her the book but not telling the loved person why he/she really comes in truth.

Intellectualization
I transform my lived experience into imaginary [SSK97].

Intellectualization is a defense mechanism by which feelings and emotions are reduced to logic. An example is a person who is terminal ill e.g. has cancer. So he/she learns about the disease and distances himself/herself from the situation that means he/she represses the illness and thinks about it only theoretically.

Regression
I turn over, I go back [SSK97].

Regression is the mental shying to a previous level of the psychological evolution that means applying pattern of behavior used earlier in development. For example in stressful situations he/she is eating or smoking excessively.

Reaction Formation
I prefer on side of my experience over the other [SSK97].

Reaction Formation is a specific and general power of the Ego, by which one of the opposed feelings is drown out so that only one is found in awareness while the other one - the wish - is shift to the unconsciousness. For instance a person, who has sadistic feelings, tries to compensate this with an extremely love to animals or other people. This person is outraged if someone tells him/her that he/she is annoyed by this extreme love.

Isolation
I isolate one felling (or concept) from others [SSK97].

Isolation means the detachment of emotional impulses from situations which causes anxiety. For example a medical student dissects a cadaver without being disturbed by thoughts of death. You see that the student isolate his work from the thought of death which can evoke a fearful feeling about his/her life or rather the fact that life has an end.

Undoing
I just take it away [SSK97].

Undoing means that a person does not want to believe that certain thoughts, words, gesture or other things did ever happen. Therefore he/she uses cogitations, words, gesture or other acts with the opposite meaning. For example someone excessively praises a person who was insulted by him/her before.

Introjection, Identification
I take it in the mouth, I absorb it [SSK97].

The Ego uses this defense mechanism to absorb object or its character from the outer in the inner world (Introjection) to become the own character (Identification). For instance a boy, who is reprimanded by his teacher, makes faces which look like an angry person. This reaction on the blame is the identification with the dreaded object of the outer world.

Turning against the Self
I beat myself [SSK97].

Turning against the Self is a defense mechanism by which a person, who has bad feelings or emotions against another person, turns them against himself/herself. For instance a child is angry about his/her father, because he has no time for him/her but the child cannot blame him. So the child becomes angry about himself/herself.

Reversal
I turn my back to these things [SSK97].

Reversal is a mechanism, which transforms the goal of the drive with the crossover from the activity to the passivity or opposite. For example a child who has been beaten from his/her parents lays hand on his/her sibling, i.e. the victim becomes the actor.

Conversion
I express something in body language to not be conscious about this [SSK97].

Conversion means the expression of inner conflicts in a somatic way. An example is a 'hysterical not be able to see' which means a 'not want to see' and refer to a desire of voyeur - 'You want to see forbidden things, therefore you shall not be able to see'.

In Table 3.1 and 3.2 different defense mechanisms are specified and an overview of their triggers and effects is given as well as examples of the human world are described. Because of not needed to divide them into different categories like in psychoanalysis, the list is not ordered.

Defense mechanism	Trigger	Effect/Impact	Example (people)
Denial	Fear, sense of inferiority, insecurity	Denial the real world, deny that the event ever occurred, deny that a problem exist	Alcoholics deny that they have a problem
Repression	Drive is confronted to contrarian claim of the Superego, i.e. like confronted to dislike	Keep information out of conscious awareness	
Displacement	For example frustration, feelings and impulses	Taking out this feelings and impulses on people or objects that are less threatening and have nothing to do with them	Frustration at work which cannot be discussed with the boss effect anger at home
Sublimation	Unaccepted impulses e.g. aggressive drives	Convert this behavior into a more acceptable form	To take aggressive drives out on sport like boxing; intellectual activities satisfy like sexiness
Projection	unaccepted qualities or feelings	ascribing them to other people	Someone does not like a person but thinks that he/she does not like him/her
Intellectualization	Feelings, emotions	reduce of feelings and emotions to logic	If someone is terminal ill e.g. cancer he/she learns about the disease and distances him/herself from the situation
Rationalization	Unacceptable feeling or behavior	try to reason why you feel or deal/act so	Student who blames a poor exam score on the instructor rather than his/her lack of preparation or knowledge
Regression	Stressful situation	Apply pattern of behavior used in earlier stage of development	Eating or smoking excessively
Reaction Formation	Anxiety	Taking up the opposite feelings, impulse or behavior	Treating someone you strongly dislike in an excessively friendly manner in order to hide your true feelings; disgust

Table 3.1: Overview of Defense Mechanisms [Fre84],[SSK97] - Part I

Defense mechanism	Trigger	Effect/Impact	Example (people)
Isolation	Unacceptable feeling or anxiety	Splitting-off of the emotional components from the thought	A medical student dissects a cadaver without being disturbed by thoughts of death
Undoing	Certain cogitations, words, gestic or other things	Substitute them with cogitations, words, gestic or other acts with the opposite meaning	Excessively praise the person you insult before
Introjection, Identification	Jealousy that someone can do things better or admire the character of someone or fear	Absorb the object or its character and they become the own one	A boy, who is reprimanded by his teacher, makes faces, which look like an angry person. This reaction on the blame is the identification with the dreaded object of the outer world.
Turning Against the Self	unaccepted feelings and emotions	turn feelings and emotions against oneself	Someone is angry about his/her parents and wants to lay into them but he/she is afraid and begins to lay into himself/herself.
Reversal	A bad experience	Transforming the goal of the drive with the crossover from the activity to the passivity or opposite	Child who was beaten from parents lays a hand on his/her child
Conversion	Dramatic situations or problems	transformation in somatic symptoms	A person, who is a voyeur, suddenly cannot see - 'You want to see forbidden things, therefore you shall not be able to see'

Table 3.2: Overview of Defense Mechanisms [Fre84],[SSK97] - Part II

Motives of Using Defense Mechanisms in the Human Mind

There exist various motives, why the Ego has to defend itself against drives. A. Freud had described in [Fre84] the following three in detail, but she noted that in the later life other reasons arise too. They emerge due to the claim of the Ego that the impulses inside have to be in agreement.

The oldest and efficient explored defense situation is the neurosis of an adult. The drive is conscious and would like to achieve satisfaction. The Ego is not antipathetic but the Superego appeals. So the Ego complies to the higher instance before a conflict between it and the Superego starts. Therefore it begins the fight against the drive. This is called the defense against the drive under the pression of the Superego fear. In this situation the Superego is a trouble maker, because it inhibits a friendly agreement between the Ego and the drive. It represents the proscribing of the sexuality and explains the aggression for antisocially. The force of the Ego is so extreme that it is not compatible for the psychologic health and makes the Ego drive adversarial and unable to enjoy. In Figure 3.2(a) and 3.2(b) the cycle of the defense of the Superego against the drive is shown.

At the beginning
1) Id to Ego: "I would like to do ..."
2) Ego to Id: "Why not, it sounds not
3) Superego to Ego: "What? Not dangerous, are you stupid! It is forbidden to do so!"

(a) Drive Defense of the Superego - At the beginning

After the influence of the Superego
1) Id to Ego: "I would like to do ..."
2) Ego to Id: "No, we cannot do that, because..."
3) Superego to Ego: "Exactly, that's ok."

(b) Drive Defense of the Superego - After the Influence of the Superego

Figure 3.2: Results of the First Test Condition

Another situation why drive defense emerges, is the real-fear in the infantile neurosis. A child, who would not like to come into a conflict with its parents and their prohibitions, also struggles against drive impulses. In the same way the Ego of a child is not against the drive, but the satisfaction is forbidden from the educator and has penalty as a consequence. As it can be in the neurosis of an adult, a child shows the same symptoms. Thus it is shown that the outer world like parents or relatives have also much disposal of the child.

Drive defense can also occur as a result of the fear of drive force. The Ego is drive friendly as long as it varies not too much from the Id. If the development from the lust-principle to the reality-principle is fulfilled, the Ego becomes drive different. The disturbance against the drive is always existent, but it is hidden from the Superego and the outer world fighting against the drive. However, if the Ego feels that it has lost the protection of both - Superego and the outer world, the enmity enhances a fear and therefore defense mechanisms are used. This phenomenon can be observed, when the power of a drive upsets the balance of the psychologic instance, e.g. in the puberty.

It can be seen that there are various motives, why human uses defense mechanisms. The defense due to the Superego and the outer world (parents, teacher, etc.) have one point in common. Both have primary lust, if the drive becomes satisfaction, but secondary unlust occurs because of

feeling guilty. This arises as result of the penalty and is the reason for defense of drive to avoid unlust. So if psychoanalysis is used for building autonomous agents, defense mechanisms has to be included in the model as well and thus this is more explained in detail in the next paragraph.

Reasons of Using Defense Mechanisms in Autonomous Agents

There are various reasons, why using defense mechanisms in autonomous agents is meaningful. One is because of utilizing approaches of psychoanalysis to explain the human mind. Another one is as a result of possibility that sensors or actuator can be defect. These motives and some others are explained in detail below. The reasons why a specified defense mechanism is used at a certain time is not known in psychoanalysis and so it is not possible to create rules when which mechanism should be applied. This has to be review for the application in which they are implemented.

The approaches of psychoanalysis includes the model of S. Freud, which uses emotions, drives, etc. This enables agents to be more intelligent as before, but other problems occur. If for instance two contrary drives would like to become satisfy at the same time, only one can be fulfilled and the other one causes frustration. To avoid this situation defense mechanisms are used to solve this conflict. Another example is the imbalance of an emotion e.g. rage, if the level of this emotion is too high, a defense mechanism is utilized to discharge this tension. Defense mechanisms can also be useful if the Superego inhibits occurred drives or desires. Because if the satisfaction of them is not allowed, frustration would emerge.

Another example for using defense mechanisms would be the one of competing goals.
Assuming that there are two desires:

- The first one, named desire A, is a desire, which could be achieved.
- The second one, named desire B, is a preferable desire, but it is impossible to satisfy it at the moment.

Because desire B is preferable, it has the highest priority. As not be able to fulfill it, frustration comes up. Furthermore the satisfaction of the desire A is inhibited, which could by chance enable desire B if it is achieved. To solve this problem defense mechanisms can be used in various ways. In the following some possibilities for solution are described:

- By using the defense mechanism Sublimation another desire, which is satisfiable, is used to reduce frustration.
- By using the defense mechanism Intellectualization or Rationalization a reason is given, why it is not meaningful to fulfill the desire B and thus reducing the priority of this desire to be able to satisfy the desire A.
- By using the defense mechanism Repression, the desire B becomes prevent, i.e. is unconscious, and so desire A can be satisfied.

This can be used e.g. as new prioritization method, because the application of a defense mechanism gives the possibility to change the values of the priority to enable other actions. It also can be a solution if a program is in a deadlock, because a process, which waits for an utilized resource

to be able to compute, can be started by repressing that the resource is used by another process. As well it can be interesting to use defense mechanisms for stopping a looped process - especially endless loops - e.g. by repression that the loop should be run again or by Intellectualization giving a reason why other processes are more important to fulfill.

Another application field could be the handling of failures in the controlling of sensors and actuators of an agent. Agents need sensors to get informations about the environment and actuators to interact with each other or with the environment. Thus a problem can occur if e.g. a sensor is defect, but is important for calculation, and can destroy the agent or be dangerous for its user. In this case, repression of the defect sensor can be done. The agent does not realize this and believe everything is ok, but the user can react without danger to this and is able to solve this e.g. by changing the sensor.

The defense mechanism Regression is used if patterns of behavior acquired earlier in the development are applied instead of new learned ones. This can be useful in a technical system if e.g. a new learned function may be dangerous in a situation where it is known that the old one always helps to avoid possible injury. So this module gives the agent the possibility to fall back to older behavior in dangerous situations. As well it can be applied by testing a new function of a module in the mind whereas the influence on the whole system can not be estimate due to the complexity. Regression can be used therefor to fall back to the old system and does not use the new function because of anxiety. This can be helpful to avoid the destruction of the system or a part of it [DTM+09].

3.2 Two Cognitive Architectures

In this Section two cognitive architectures of the project ARS are described. The first model is used as base for the first prototype - the Bubble Family Game (BFG). As well the technical implementation of the defense mechanisms is done in this model. The second model has been developed, but it is not implemented in a first simulator at the moment. That's the reason, why the implementation of the defense mechanisms is not realized in this model, but it is explained in the Section 3.2.2, to which parts of the model they have connections, how the flow of information happens and in which part it is situated.

3.2.1 The First Model

A new cognitive architecture for automation systems and autonomous agents is developed in the project ARS (Artificial Recognition System). This comprehensive architecture combines low-level, behavior-oriented forms of decision making with higher-level, more reason-oriented form of behavior generation and selection. This model is based on psychoanalytical theory - the second topographical model or also named structural model of S. Freud, which will be complemented with established findings from neurology, neurobiology, ethology and similar sciences, as long as there are no contradictions. This is necessary, because the structural model tries to analyze a very complex but already existing structure top-down and naturally lack a lot of information about details, which have to be specified when constructing and synthesizing intelligent system bottom up. It has been proven now that neuroscientific findings can validate some of Freud's most central assumption, thus narrows the gap between neuroscience and psychoanalysis considerably. The technical design includes both directions - bottom-up as well as top-down - in order to keep white

spots in the cognitive architecture as small as possible. However, there is no intention to copy the brain on structure level but to create an artificial system that aims the functional equivalence. In the technical design, there will be incorporated lots of other psychoanalytical principles in addition to the chosen basic psychoanalytical model, shaping structure as well as processes of architecture. In the following, the neuro-psychoanalytically inspired cognitive architecture will be presented in detail [PPC09], [Pal07].

Inner and Outer World

According to [Sol04] the inner world constitutes to the mind and not to the bodily states. A fundamental aspect of this model is, that the mind is anchored in a body and two kinds of signals have to be processed - internal (bodily) and external (world). Outside world as well as internal states are influenced by actions. This concept is compatible with the psychoanalysis as well as neuroscience. Various neuroscientists have stressed that our subjective, psychological inner world is inherently anchored in monitoring our body - the way we feel ourselves. In their theories, the signals from self, which is represented by bodily needs and regulatory elements (value system), is linked with the signals from the surrounding outer world and this is critical for consciousness obviously making organisms more intelligent. Important is that the inner world is the mind and the outer world is the body as well as the surrounding environment. So integration of external affordances, internal needs and appropriate actions can be viewed as a key element of intelligent behavior, especially if integration is connected with learning and an evolutionary memory that goes beyond 'the current present' [PPC09].

Interaction and Feedback

In recent years, it has been realized that actions are guided by perception - in fact the origin of intelligence. Organisms cannot just lay back and take their time, but they have to fulfill their bodily needs. What kind of actions an organism can perform is prescribed by the particular form of organisms embodiment. Furthermore, there is always feedback from the environment, which is again perceived, and enables the organism to interact with the environment. The feedback has the form of sensorimotor or emotional experiences, which serves as the basis for development of cognitive categories and concepts. The simplest case are just hard-wired rules between perception and action. They can be viewed as external regularities having been internalized during the course of evolution, but the incorporated process between perception and action has become more and more complex. For instance, solely unconditioned reactions have been available, by opening simple control loops, conditioned reactions were implemented which possess a higher degree of flexibility and context-sensitivity. Members of a list that features distinguishing levels of homeostatic control are mechanisms like reflexes, drives, instincts, conditional associations or emotions, which have been acquired during evolution. More 'intellectual sorts of thinking' are for example reasoning, drawing inference or planning. It is clear that better kinds of representations carry an improvement in homeostatical control. So far, no one is able to tell exactly how complex relationships, especially symbolic ones, are represented in the brain. It is known that various kinds of memory are possessed by human, whose structure get gradually functional and more and more 'filled' with contents during the lifetime. However, remembering explained in detail in [Gru07] is not just an act of passive retrieval of some fixed set of data, but is an active construction process - a great deal of which happens unconsciously [PPC09].

General Architectural Scheme

Figure 3.3 shows the cognitive architecture, technical implementing the neuro-psychoanalytic view of the human mind. The architecture consists of two main blocks named Pre-Decision and Decision unit. Each of these blocks contains various functional models, which again houses further sub-structures and processes and there exist different kinds of memory systems. The various functional modules and the memory units are connected by information and/or control flows. However, the modules of the architecture certainly do not anatomically correspond to the structures in brain. Also most functions bundled within a module are probably in reality much more 'spread out' over the brain and much more intertwined with each other. The present architecture shown in Figure 3.3 is a first attempt of a technical implementation, a feasible scheme, that shall be further improved iteratively.

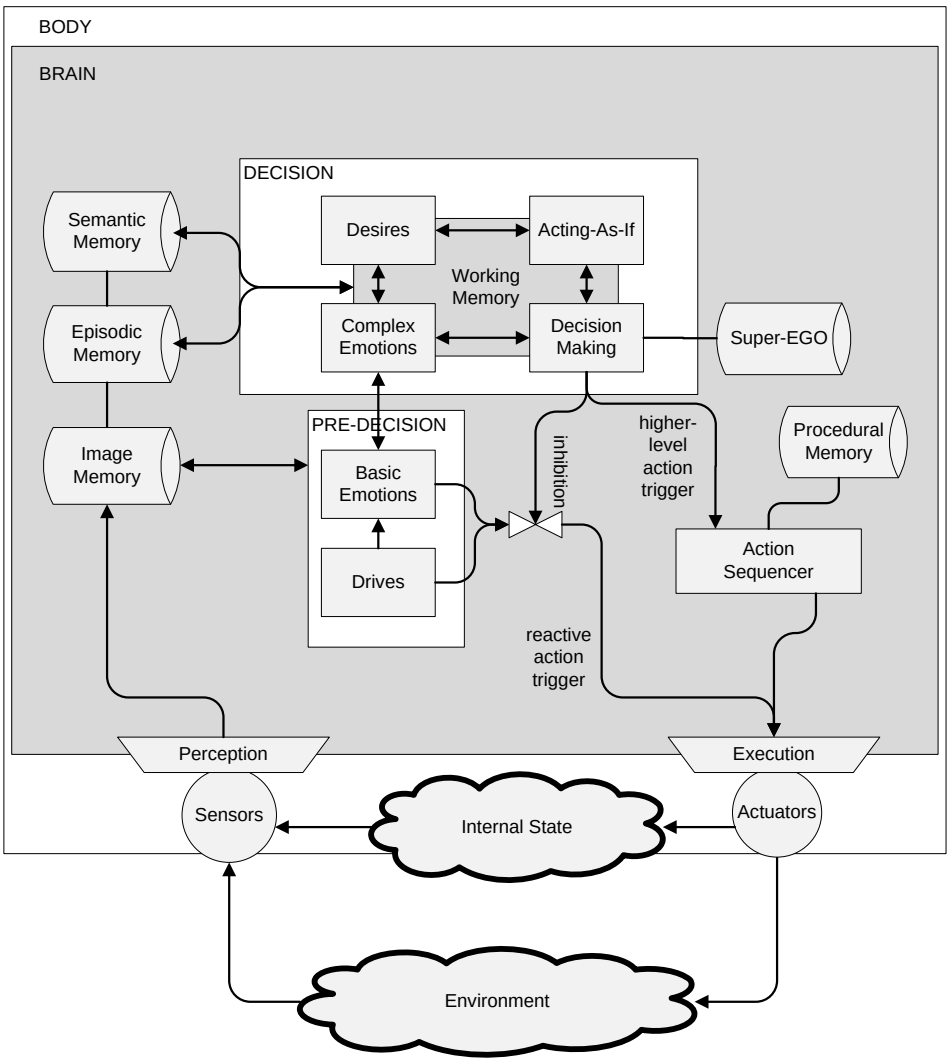


Figure 3.3: Model Used in ARS-PA Project

The key ideas of the neuro-psychoanalytic picture that have been implemented are:

1. The 'body' of the autonomous system and the environment as elements are integrated in the form of a feedback loop.

2. The combination of low-level and high-level mechanisms are summarized in two main functional blocks: the Pre-Decision unit, corresponding to the psychoanalytic Id, and the Decision unit, corresponding to the psychoanalytic Ego.
3. Emotions are used as important integrative and evaluative elements, in combination with drives and desires.

Freud did not use the term emotions but applied affects. He saw them as subjectively experienced manifestations of underlying physiological changes. He connected them with his concept of drives, which originate from 'within the body' in response to 'demands placed on the mind in consequence of its connection with the body'. Technically speaking, information processesing systems are enabled by emotions to learn values along with informations they acquired. Therefore, the introduction of an episodic memory, which contains emotionally evaluated previous experiences, as an element of the architecture is very important [PPC09].

External and Internal Perception

The *Perception* module contains the filtering and symbolization processes of internal and external sensor data. Both are continuously streaming into the system, which causes an enormous amount of information. This amount of information has to be reduced by filtering the relevant pieces of information. The result is the perception of external objects and situations as well as the perception of internal bodily states. The symbolization process of the external perception consists of several levels, which are hierarchically organized. During childhood, our episodic and semantic memories are gradually built up, and the consequence of this is, that the perceptual process is more and more governed by deeply encoded and abstract knowledge derived from learning experiences. So human perception is shaped by previous memories and not unbiased and neutral. Internal stimuli are received in the *Perception* module parallel to external stimuli and the part of the *Perception* module, which is dedicated to the internal perception, watches over the 'bodily' needs of the autonomous system. These needs are represented by internal variables and each of them manages an essential resource of the autonomous system that has to be kept within a certain range [PPC09].

Memory Systems

A central feature of the presented neuro-psychoanalytical approach is, to shape, and also to speed up cognitive processes by using of predefined 'mental images' as templates. Thus, perception is not possible without memories, which is actually a process of re-combining current inputs with previous experience. So the role of memories is to organize previous experience into recognizable chunks of knowledge. Depending on the cognitive level, the current goal or task and the current stage of the decision making process, different types of memories are involved. The *Image Memory* is the most basic form and is mainly associated with the perception process. It contains representation of symbols, objects and situations and is a kind of passive, partly implicit, short term memory. The *Working Memory* is part of the decision unit. It is an active, explicit kind of short-term memory and supports the higher-level cognitive operation by holding goal-specific information and streamlining the information flow to the cognitive process. The *Semantic Memory* includes facts and rules about the world. Examples therefor are physical rules of the environment or how objects relate to each other. The content of the *Episodical Memory* in

detail explained in [Gru07] is based on autobiographic events, which are remembered from a first-person view. The *Superego*, which is viewed as a special part of the *Episodical Memory*, includes the rules for socially acceptable behavior. The *Procedural Memory* is a kind of bodily memory, which holds necessary informations for the execution of behavior, for example the knowledge how to run [PPC09].

Low-level Decision Making - Pre-Decision Unit

Freud expressed, that most processes, which determine our everyday feelings, wishes and thoughts, occur unconsciously. However, the individual's well-being is ensured by unconsciously motivated activities of the brain, which are not random and so drives and affects play an important role. Freud conceptualized drives as a signal of the needs to the mind of the body. Affects, which possess physiological as well as psychic aspects, were ascribed a similar intermediate role. The unconscious, instinctual forces are enclosed in the Id, which is governed by the 'pleasure principle'. Today, in the model of J. Panksepp [Pan98] the following four basic emotional systems exist, which are influencing the individual and its acts:

SEEKING-System
: This system has a special functionality, as it has the destination to assure the balance of the internal states of an organism. Therefore SEEKING-System creates excitation and energy which inspires our interest in our environmental world and activates the searching for the object which satisfies the needs. A subsystem is the LUST-System, which shows the brain, that the object satisfying the need is found, and is only activated when the SEEKING-System is deactivated. So The SEEKING/LUST-System is considered as a pleasure-seeking system, which motivates most of our goal-directed interactions with the world. Pleasure creates an urge of repetition of the pleasurable behavior, whereas frustration leads to avoidance.

RAGE-System
: This system is activated when frustration and restraint or body irritation is appeared.

FEAR-System
: This system ensures that an individual can escape rapidly from dangerous situations respectively avoid dangerous situations. Two different reaction are able to be caused. The first one is a congealment-reaction. A probably reason therefor is, that an object which moves is easier to detect as an object which stands quietly, but only be expedient when the enemy is far away. The second one is the escape.

PANIC-System
: This system is associated with panic as well as with the feelings sorrow or loss and if it is stimulated, the seeking-behavior and the vocalization are activated. Such behavior increases the possibility to be found by someone or to find someone.

The RAGE-System, the FEAR-System and the PANIC-System, but not the SEEKING-System, are mainly influenced by sensors of the outer world [Sol04],[BLPV07].

High-level Decision Making - Decision Unit

If the *Pre-Decision* unit has not already 'fired', the more exhausting evaluation and decision making processes of the *Decision* unit are running. The central task of the *Decision* unit is the

association of desires, (basic and complex) emotions and thoughts in such a way that a release of an 'intelligent' action command follows. Mental life has been described by Freud as a kind of continuous battle between conflicting psychological forces such as wishes, fears and intentions and the resolution of the conflicts is the compromise among competing motives. In the mental process occurs the great deal unconsciously and the conscious part of the mind has only limited access to the unconscious parts. In this respect, it appears that the role of emotions is to inform the higher cognitive apparatus how world events relate to intrinsic needs. The *Decision* unit also has the possibility to inhibit the execution of the impulses caused by the *Pre-Decision* unit but not to prevent their arousal. More complex behavioral pattern can also substitute original, primitive action tendency. The desires of the model have a similar nature as the drives and are also the motivational source of actions, but now, more sophisticated and more enduring actions. Desires enables behavior that aims to improve the subject-world relationship and a fulfilled desire discharge one or more tensions. The wish to re-experience a once satisfying situation is named desire and its data structure has to involve a need, an object of desire and a representation of the self and its expectations regarding the interaction with the object and the fulfillment of the desire. The *Complex Emotions* module contains all emotional mechanisms that go beyond the simple ones included in the *Basic Emotions* module, which have been classified by Charlotte Roesener [Roe07]. These mechanisms deal with derivations of the basic emotions, which are labeled 'complex emotions', but also with the basic emotions. This means that a basic emotion like anger can not only be elicited or modified at the level of *Pre-Decision* unit, but also by the appraisals at the level of the *Decision* unit. Examples of task, to which emotions contribute, are to inform the self about important changes in the body and in the environment or to support the pursuit of goals as well as their abortion in time in case of turning out to be destructive. The *Acting-As-If* module creates 'thoughts' and as the name indicates is thinking viewed as a kind of acting, but there is no execution of the invoked motor commands. Scripts for typical situations, which are given by the episodic memories, are used to expand desires or tasks into goals and sup-goals. They are also utilized to evaluate the expected consequences of the planned actions, which is supported by knowledge from the *Semantic Memory*. Another task of the *Acting-As-If* module is to constantly experiment with actions, their costs and benefits in an anticipatory way. This should help to speed up cognitive reactions by narrowing the focus of attention [PPC09].

In this model emotions have influences on low-level decisions as well as on high-level decisions. In Section 2.1 the first approaches of using emotions in autonomous agents are explained. The difference between this model and the model of Canamero, which is described in detail in Section 2.1.1, is that he only implements a set of basic emotions and the model incorporates only low-level mechanisms. The appraisal based models, which are presented in Section 2.1.3, also vary to the neuro-psychoanalytically-inspired model, because the classification of the emotions is done with the dependence of the stimuli, which causes the emotion. In the model of the project ARS, the division of the emotions is done in basic and complex emotions. The difference between them is that the basic emotions are found in the human brain but also in the mammalian brain. If the complexity of the situations, in which the individual is involved in, increases, the emotional evolution is getting an emotional spectrum with a broader bandwidth, which includes learned experiences. So complex emotions are defined as an emotional system, which is not evident in mammals soon after the birth. But they are also a cognitive rule system, which is highly influenced by social rules [BLPV07]. As well no feedback from the emotional system to the cognitive system is implemented in the appraisal based models. But both, the appraisal based models and the first prototype of the project ARS - the Bubble Family Game presented in detail in Section 2.2, are a knowledge-based system, why learning is not implemented. The model explained in Section 2.1.2

is a first try to use learning in an autonomous system. However the first implementations of this model by J. Velasquez also remains low-level, because the focus is only on basic control circuits, which are implemented via connectionist networks, and symbolic representations are not explicit used.

Actions and Behavior

Actions influences external circumstances and internal states and are built with action primitives. *The Action Sequencer module* is responsible for handling sequences of actions, which can be loaded from the *Procedural Memory*, where they are stored. When they are occurring repeatedly, *the Action Sequencer module* discovers new action pattern and transfer them to the *Procedural Memory*. They can be activated in future as a whole from the *Procedural Memory* and forming consequently new kinds of routine behavior [PPC09].

3.2.2 The Second Model

The first model of the project ARS explained in detail in Section 3.2.1 raises several questions in its design. It is partly developed using different theories, which are sometimes incompatible. This happens because of the various models are described highly detailed in some areas, but other topics are incomplete or imprecise. The new model of the project ARS described in detail in [DFZB08] is based on the second topographical model of S. Freud only, which is presented in Section 3.1.1. Parts with gaps or imprecision are marked for further work, which is done when the model based on Freud is completed. Therefore engineers and psychoanalysts try to find a way to translate Freuds definitions and explanations into a technical language.

Functional Description of the Modules for the Three Topmost Levels

In Figure 3.4 the new model of the project ARS is shown. The *Brain* module forms the first layer. One layer beneath is the functional devision of the *Brain* module into *Sensor interface*, *Actuator interface* and *Psychic apparatus*. The *Sensor interface* module is responsible for the collection of information from the environment by the sensor system of the body. The *Actuator interface* module is the link between the *Psychic apparatus* and the mechanical body elements. The *Psychic apparatus* module, which is the central part of the model, creates resulting control decisions and the incoming data of the *Sensor interface* is handed over to this module. The *Psychic apparatus* module is devided in the sub modules *Ego*, *Id* and *Superego*, which fit together with the second topographical model. The arrows notify the information flow within the *Psychic apparatus* module and to its environment. The focus of communication is in the *Ego*. To offer the possibility for communication, every module has to provide an interface therefor.

The *Id* presents the source of the mental apparatus as well as the source of psychic energy and upcoming drives. The *Id* is the only part of the *Psychic apparatus*, which exists after the birth. The *Ego* and the *Superego* emerges out of it. The *Id* is responsible for the managing of the drives and its demands and forms the organic processes into psychic processes by use of representations that are called affects and presentations. The drive evolves out of sources and becomes discharged by achieving an aim by the use of the so named drive object, which causes satisfaction. It notifies the bodily needs and discharging means a reduction of them. The processes of the *Id*, which are named primary processes, are unorganized and contradictory.

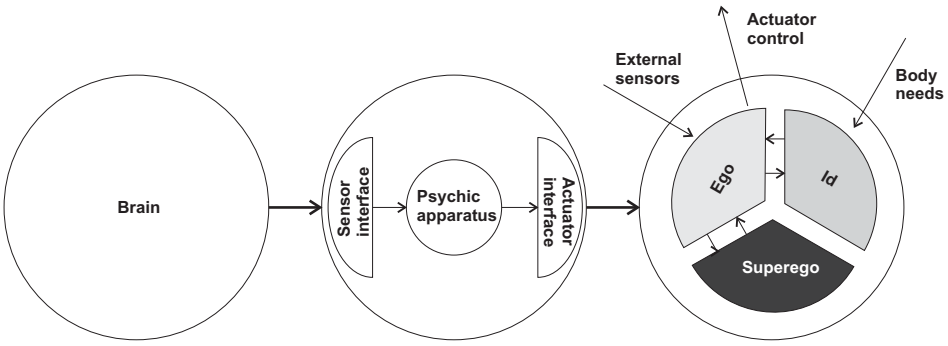


Figure 3.4: New ARS Model [DFZB08, p. 56]

Numerous conflicting, unstructured processes appear in the human mind and those, which achieve the *Ego* to be considered into the decision process, become organized. Hence they are arranged by the secondary process. If a process, which lowers the drive tension cannot be carried out, the mind tries to redirect this discharge through other processes. But therefor functions of the *Ego* are a prerequirement. The functional model of the *Id* is shown in Figure 3.5. The *Physiological to psychic transformation* module handles the transformation of physiological stimuli into psychic representatives - affects and presentations. This transformation has as representation the drive, which is used as the input of the *Quantification of drives* module and describes the interconnection between physiological and psychic processes. The *Quantification of drives* module is responsible for allocation of psychic energy. Hence an evaluation has to be done and subsequently it is tried to refer the drive representatives to the *Ego*. However the task of arranging and structuring these processes is not part of this module, but is done by the *Ego*. The *Selection of memory traces* module is responsible for the generation of memory traces and their cathexis.

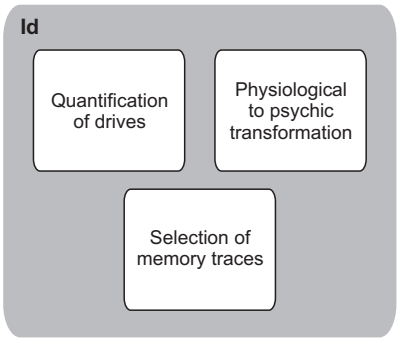


Figure 3.5: The Functional Model of the *Id* [DFZB08, p. 58]

The *Superego* evolves in the early childhood and has its roots in the *Ego* and the *Id*. Children are faced to a large number of restrictions and demands by the parents, wherefore processes are developed to repress specific drive demands in order to satisfy internalized demands. A parallel process is used to form a mental structure based on parental restrictions, demands and awards. So a task of the *Superego* is to deal with all restrictions and demands. The *Superego* gives also a feedback on actions, which are done by the subject, in a positive as well as in a negative way, wherefore it can be seen as a kind of reward system. The *Superego* and the *Ego* synthesize demands of the *Id* and the reality, which has as result control, satisfaction and inhibition of

impulsive actions. So the *Superego* fits the individual's behavior and thinking to the demands of the reality, to social rules and expectancies. In Figure 3.6 the three sub modules of the *Superego* are shown. The *Management of demands* module is responsible for arranging of demands on the basis of social norms. This structure emerges by the individual's social background and allows ideal behavior and thinking. Restrictions, which are established by the social background, are arranged in the *Management of restriction* module. Addition are possible over years, but the mainly forming is happened during the childhood. The *Management of the Ego ideal* module manages the ideal image of oneself and emerges out of the identification with the parents.

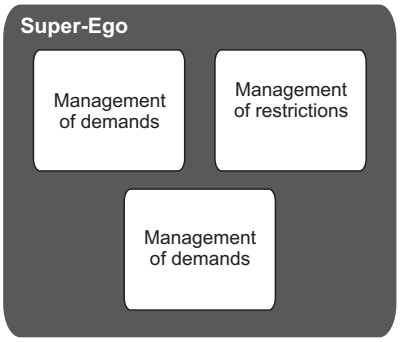


Figure 3.6: The Functional Model of the *Superego* [DFZB08, p. 60]

The *Ego* transfers information of the observed environment to the *Id* and the *Superego*. The generated images have as effect a preferred action in both, the *Id* and the *Superego*. This action is back forwarded to the *Ego*, where a comparison with the rules of the *Superego* is done. If the rules disagree with the instinctual impulses of the *Id*, a conflict is caused, which is ideal balanced by the *Ego*. Which content becomes approved depends on e.g. the relevance of information, the environment conditions and instinctual impulse. This choice is done by the *Decision management* module, which is placed in the *Ego*. So the *Ego* assumes a mediation role between the inner and outer world as well as between the *Id* and the *Superego*. Though it is important that after birth the *Psychic Apparatus* is handled by pleasure and unpleasure, because it only consists of the *Id*. By and by, the *Ego* and its functions like reality testing and focus of attention are formed out of the *Id*. In Figure 3.7 the two sub modules of the Ego are shown. The *Psychic synthesis* module picks up information of the *Id* and the *Superego*, whereas the *Id* provides informations about current drive demands and the *Superego* allocates the relevant behavior rules to the *Ego*. So this module is responsible for finding an agreement between the instinctual impulses of the *Id* and the demands of the *Superego*.

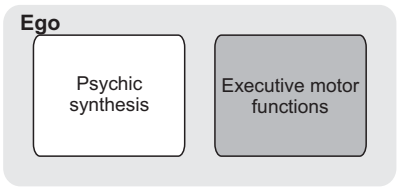


Figure 3.7: The Functional Model of the *Ego* [DFZB08, p. 61]

The sub modules of the *Psychic synthesis* module are shown in Figure 3.8. The *Perception* module acts as a pool for outer and inner world information and three modalities of perception are appointed in the psychoanalysis. The first one is an information, which does not activate a

memory trace, because no fitting memory trace exists. This occurs if the information is new to the subject and therefore has to be generated. The second one activates spontaneously a memory trace. The third and last one, which is named attentive perception, is managed in the *Focus of attention* module, which means that the looking for a needed object in the outer world, which conforms to an activated presentation due to a drive demand, is declared in this module. So the information, which is forwarded by the *Perception* module to the *Focus of attention* module, is evaluated in the latter one. This attention process can be seen as a kind of cognitive selection and concentrates the selection process on one specific type of incoming information, when this module is enabled. With the attention process an optimized information selection process should be reached, which should speed up the reaction on environmental changes. Moreover the chosen information is forwarded to the *Presentation management* module and activates memory traces, which are already available.

Corresponding to Freud, a drive has two components of representation in the *Ego*. The first one is the affect and the second one is the other representation. This affect is formed by the *Affect management* module and is sent to the *Ego* module. According to Freud, two types of presentations are available - the thing presentations and the word presentations. The thing presentations fit to the sensorial characteristics of an object and mainly belong to the *Id*. The word presentations mean the description of an object, event, etc. via a set of symbols. The organization of the facilitations between these two presentations, which is mainly formed in early childhood, is done by the *Presentation management* module.

As written before, the processes in the *Id* named primary processes are unorganized. The structuring and organizing of the content in these processes is done by thinking, whereas the organizational process is based on culturally influenced logic and has as result the conversion into the secondary process. The appraisement of the current situation as well as the comparison with sooner experienced situations is done during the process of thinking. Thus the *Thinking* module applies e.g. already existing memory traces, facilitations, etc. But also new memory traces can be evolved by this process as well as the direct execution of action can be repressed. By trying out and planning in the mind, potential dangerous situations can be prevented and the bodily energy is utilized in an economical way.

To differ between the inner presentation and the outer reality, a reality testing is done, which permits the differentiation between imagination and reality. If a failure happens in this process, the subject can be injured e.g. because of misinterpretation of a situation. This is done by the *Reality testing* module. After activation of an internal presentation by a drive demand, the attention is concentrated on the drive object. The arranging of the memory traces, which are used by the *Ego*, is done by the *Memory management* module. As well generation and activation of memory traces, representing a word presentation, are handled by this module. Observed information initiates memory traces or is redirected to the *Selection of memory traces* module, which is a sub module of the *Id* module. It is possible that a conflict occurs within the *Psychic apparatus*. This is managed by the *Decision management* module, which receives and evaluates information from the *Reality testing*, *Perception* and *Superego* modules. In the information stream of the *Reality testing* module is the received information of the *Id* module included, i.e. the *Reality testing* module receives information from the *Id* module and sends it within its information stream.

In the *Decision management* module the proper actions are analyzed and activated for the current situation by a process, whereas the generated information is forwarded to the corresponding modules. It is possible that instinctual impulses are not discharged by the decision. If this

happens, other functions of the *Ego* like the mechanisms of defense deal with them. The repression of drive discharge can be caused by the current environmental conditions or by the Superego, which is a normal process in every individual. A mechanism of defense can be explained as an inner fight on forbidden drive demands. So the *Ego* determines which thoughts becomes aware in the mind and evolves a censorship for splitting into allowed and forbidden stimuli. There exists various of mechanisms of defense and depending on the current inner and outer situations a specific mechanism is used. A mechanism of defense can repress or distort a drive demand as well as an instinctual aim can be displaced, suppressed or inhibited. The input therefor comes from the *Decision management* module.

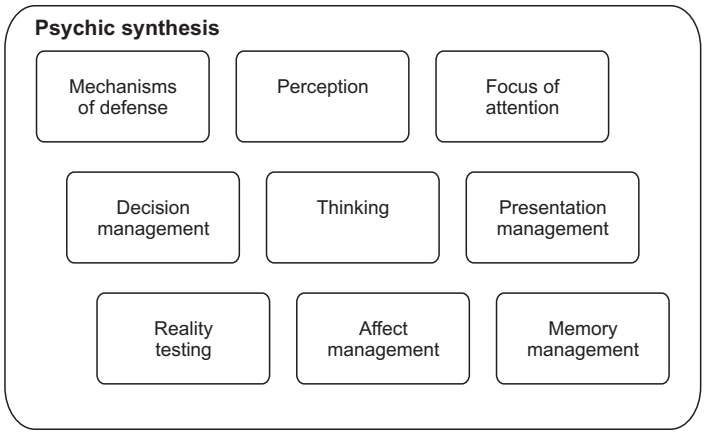


Figure 3.8: The Functional Model of the *Psychic synthesis* [DFZB08, p. 61]

In Figure 3.9 the two sub modules of the *Executive motor functions* module, which is responsible for the controlling of perception and motor activity, are shown. The *Focus of attention management* module is invoked by the *Psychic synthesis* module, which needs specific information for being executed, and decides on necessary environmental information for leading the *Focus of attention* module - a sub module of the *Psychic synthesis* module (see Figure 3.8). An example therefor is the focus of attention for locating the drive object for discharging the drive. The *Decision Management* module - sub module of the *Psychic synthesis* module - provides the *Motility control* module with informations, which are used to govern the body's motility and motor driven actions.

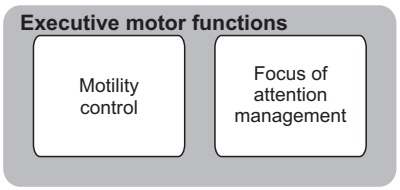


Figure 3.9: The Functional Model of the *Executive motor functions* [DFZB08, p. 63]

It is important to say that this model is not completed yet. In the current work the further developing of the various layers, the assignment of the information flow between the modules and the forming of the interface between the *Psychic apparatus* module and the environment is done. Furthermore it has to be found an interface between the new model of the project ARS and the sensor data processing theories, to provide the decision unit with the data from the outer world.

3.3 Modeling Defense Mechanisms

For technical realization a general model of the defense mechanisms has to be developed. Therefor a view on human usage of defense mechanisms is done first, to find out important characteristics of them. Afterwards these found characteristics are implemented in a general technical model. Finally, a model and possible realizations of different mechanisms for the Bubble Family Game (BFG) are described.

A General Model of Defense Mechanisms

To be able to create a general model of defense mechanisms for technical realization a detailed view on human is done. Humans experience a bad situation and are not able to deal with it or have drives, which are possible not allowed by the Superego, the outer world or are contrary. This causes an inner conflict and the tension rises. If the tension cannot be discharged directly and exceeds a threshold, a defense mechanism is used. The defense mechanism gives the possibility to discharge the inner tension in a specific way, which depends on the chosen mechanism, and to deal with the inner conflict. Important is, that humans do not realize that a defense mechanism is used, i.e. this happens unconsciously. So it is seen that a situation has to happen or an 'unsolvable' problem occurs, which causes an imbalance, before a defense mechanism is used. This imbalance can be seen as an *input*. There is also always an *output*, e.g. the defense mechanism Sublimation has an action (=*output*) like boxing for reduction of aggression. Other possible *outputs* are 'do nothing' or 'not be able to move a hand'. Both, *input* and *output*, depend on the used defense mechanism. In the following this model is transformed into a technical model for the BFG.

A Technical Model of Defense Mechanisms

Derived from the general model, the technical model of defense mechanisms, which is shown in Figure 3.10, needs an input and an output. The input is the trigger, which activates a defense mechanisms. Two different inputs are possible:

- The first one is caused by the inner world, e.g. when not satisfied drives or desires increase levels of emotion like rage and disappointment.
- The second one is caused by the outer world, e.g. a situation happens like seeing too many enemies.

The exceeding of the level of an emotion over a threshold involves a defense mechanism, which is responsible for discharging the tension. This is done with setting an action such as in Sublimation or a somatic symptom such as in Conversion. The action can be active or passive. An active action is e.g. attacking someone else or moving against an obstacle. A passive action is e.g. do nothing, although another agent needs help. Thus, in the technical model two inputs, which also can be named trigger, and two outputs, which also can be named effects, are possible.

In the BFG bubbles have basic and complex emotions, which have different levels. Various possibilities to arise the level of emotions exist. For instance this can be caused by a drive or a desire, which is not satisfied or by seeing a group of bubbles of another team, which can attack it. In the case of rising over a threshold a bubble can use a defense mechanism. It is also possible to variable the threshold value which is not independent from other emotions e.g. if a bubble has

more hope for satisfaction of its desire will use a defense mechanism later than a bubble with less hope. It is also possible that more situations have to happen to activate the defense mechanism like two or more desires have not been satisfied. An example therefor is the defense mechanism Conversion, which is explained in detail in the next chapter. To reduce the level of an emotion the defense mechanisms uses different ways. One of them is to attack or do not help another bubble. Another one is to deactivate sensors or to make them fuzzy, e.g. a bubble sees fuzzy, which means that it is not able to locate another bubble or an energy source exactly.

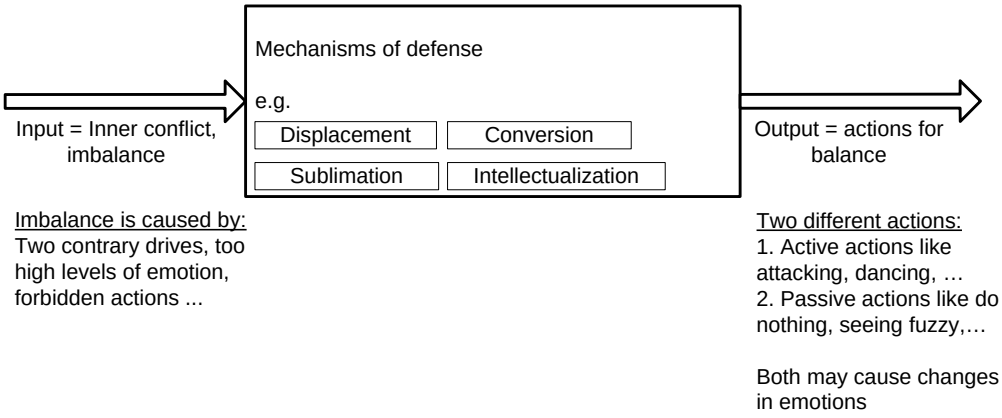


Figure 3.10: Model of Defense Mechanisms for the BFG

As explained in Section 2.1.5 Buller also implements some defense mechanisms. Though he distinguishes in [Bul05] that his approach is only based on psychodynamics and not on psychoanalysis, i.e. that the existence of the Oedipus complex and the key role of sexual drives and aggressive drives in the development of personality is not contained in the assumptions of the basic theory, which he uses for designing his model. He defines the psychodynamic perspective as follows:
"(...) that the psychodynamic perspective is a severely expurgated version of Freudian work, i.e. free of the most controversial statements about sexuality and aggression."[Bul05, p. 72]
The model of the project ARS explained in Section 3.2.1 is based on the second topographical model of S. Freud, which is described in detail in Section 3.1.1. In this model the Id is the part, in which instinctual drives occur. According to Freud, drives are a signal of the needs of the body to the mind, which would like to become satisfied [Köh93, p.143]. The source of the drives is a somatical process in an organ or in a part of the body, which is represented as drive in the psyche [Köh93, p.143]. Freud developed different drive models and attended to various drive destinies like Repression or Reversal, which had been categorized to one of the defense mechanisms by his daughter A. Freud. The model of the project ARS differs from the model of Buller by this main approach of using drives. Buller only utilizes the difference between his four realities for initiating the defense mechanisms, whereas with the difference a tension is calculated for controlling them. But in the model of the project ARS the imbalance of the satisfaction of the bodily need evokes a defense mechanism to reduce this imbalance, which has as effect the rising of negative emotions like rage or fear.

3.3.1 Integration into the Two Models

As described in Section 3.2, there exist two model of the project ARS. The defense mechanisms are implemented in the *Mechanisms of defense* module and are situated in both models. Where

this module is placed and how the information flows is explained in the next two paragraphs.

The Defense Mechanisms placed in the First Model

In Figure 3.11 the module *Mechanisms of Defense* is embedded in the first model of the project ARS, which is explained in Section 3.2.1. As seen in Figure 3.11, the module is situated in the Decision unit, because it is part of the High-level decision. The module has various inputs and outputs, which are demonstrated by the arrows. The following three inputs are possible:

1. The first input are the levels of the basic and complex emotions shown by the arrows from the *Basic Emotion* module and the *Complex Emotion* module to the *Mechanisms of defense* module.
2. The second input are the levels of the drives shown by the arrow from the *Drive* module to the *Mechanisms of defense* module. This input is necessary because of not being able to discharge the drive, a defense mechanism is activated to restore the balance.
3. Because of influencing the *Decision making* module by *Superego*, which can forbid actions, the *Decision making* module also has the possibility to evoke a defense mechanisms e.g. to find an alternative way for discharging a drive (Sublimation). As well it is possible that two contrarian drives would like to become satisfied at the same time. Because of not be feasible, the *Decision making* module has to be able e.g. to repress one drive, which means this module must be able to activate a defense mechanism. This is the third input, which is shown by the arrow from the *Decision making* module to the *Mechanisms of defense* module.

As known from the general model of defense mechanisms, they also generate an output. The following two outputs are possible:

1. The defense mechanisms are activated because some levels of emotions are imbalanced. So they must have the possibility to change these levels e.g. the defense mechanism Displacement is used for reducing the level of rage. This influence on levels of emotion is shown by the arrows from the *Mechanisms of defense* module to the *Basic Emotion* module and to the *Complex Emotion* module.
2. To be able to reduce levels of emotions, it must be feasible to set actions e.g. for the defense mechanisms Displacement the action 'hit someone, who has nothing to do with the problem'. Thus the *Mechanisms of defense* module has to be able to forward the needed action to the *Decision making* module. This is shown by the arrow from the *Mechanisms of defense* module to the *Decision making* module.

The Defense Mechanisms placed in the Second Model

As explained in Section 3.2.2 the *Mechanisms of defense* module is embedded in the *Psychic synthesis* module, which is a submodule of the *Ego* module. By the *Decision management* module it is decided, which desire becomes satisfied and which not. If a desire cannot be fulfilled an inner tension rises, which causes an imbalance. Because of not being able to reduce this tension in a direct way the *Mechanisms of defense* module becomes activated. In this module the decision, which mechanism should be used to discharge the tension, is made. The chosen mechanism is

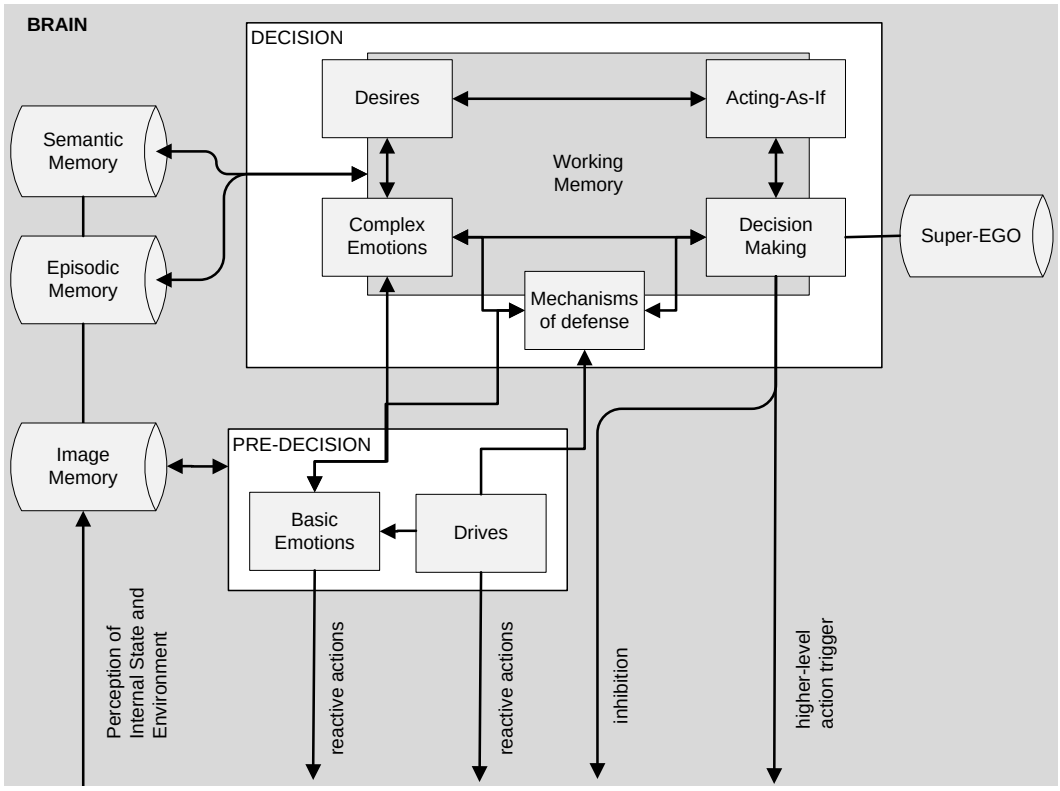


Figure 3.11: The Module *Mechanisms of Defense* Embedded in the First Model of the Project ARS

transfered to the *Thinking* module to check, if it is possible to utilize it. If it is not appropriated to discharge the tension in that way, the *Thinking* module transfer this information back to the *Mechanisms of defense* module. Otherwise all information is forwarded to the *Decision management* module and the actions, which have to be done for applying the mechanism, achieve the highest priority and are executed as soon as possible. If no actions, which are executed in the outer world, are necessary the information is not transfered from the *Mechanisms of defense* module to the *Thinking* module but directly to the *Decision management* module to become carry out as soon as possible. If it is not possible to use the chosen mechanism, the *Mechanisms of defense* module searches for another one and forwards the information again to the *Thinking* module if necessary. The information flow is shown in Figure 3.12. The involved module blocks are bordered with a double line and the information blocks, which is forwarded to the modules, are bordered with a thick line.

3.3.2 Characteristics and Influence on the Behavior

In this Section some examples of possible realization of various defense mechanisms in the BFG are described. As well it is analyzed what is necessary to be implemented in a model, which uses approaches of psychoanalysis. All defense mechanisms, which are explained in this chapter, are only ideas of possible realizations and are not implemented in the BFG. For the four defense mechanisms Displacement, Sublimation, Intellectualization and Conversion a general model is developed and explained in Section 3.3.4. The technical realization of them is presented in the next chapter in detail.

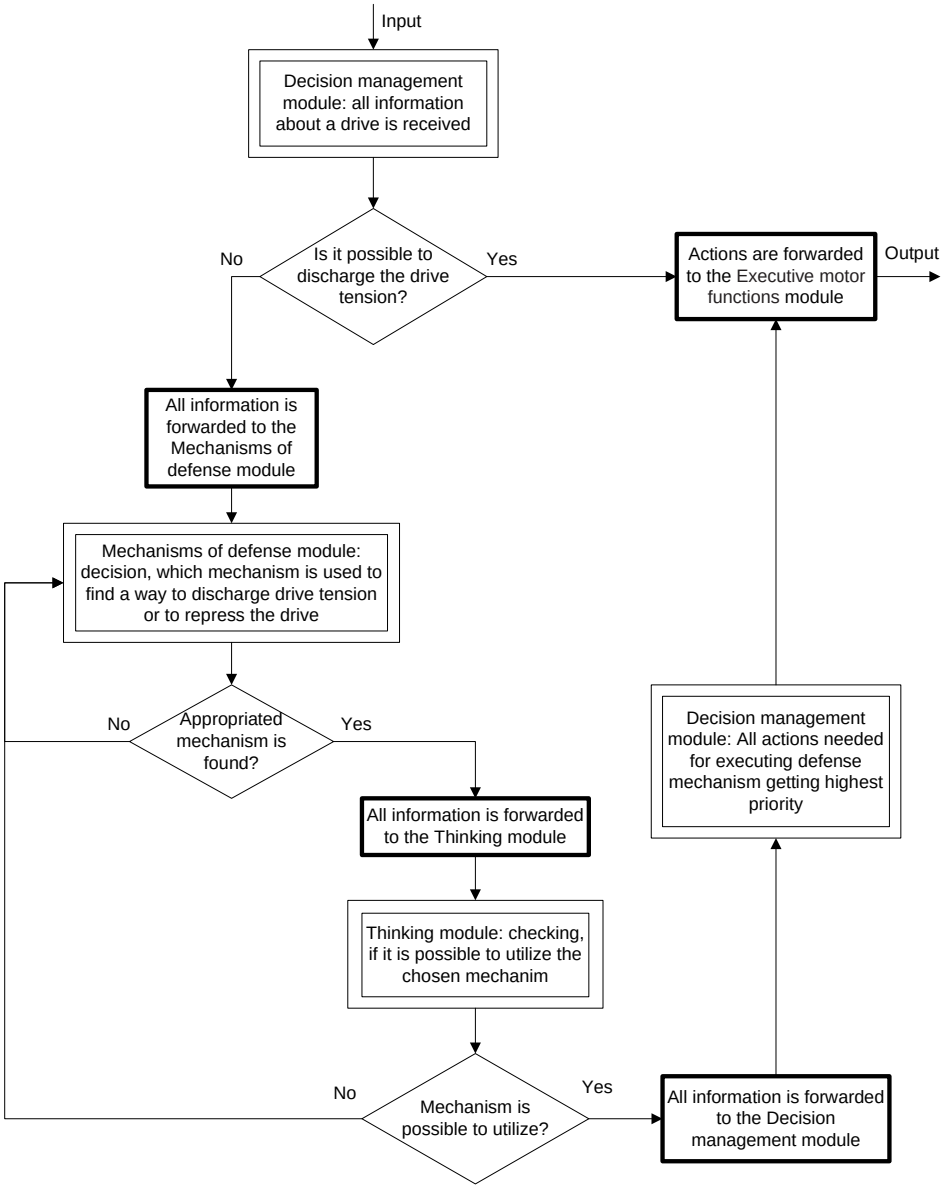


Figure 3.12: Information Flow of the Application of the Defense Mechanisms in the Second Model of the project ARS

Repression

This defense mechanism is used to repress e.g. phantasy, happened situations or one of two contrary drives/desires, i.e. they should not become conscious. An example for the BFG is a bubble, which is hungry, because its energy level is very low, and it sees an energy source to satisfy the desire to eat. But it also has detected many enemies, e.g. many bubbles of the opposing team, and is afraid of being attacked by them. So it has the desire to flee. Now two possibilities exist. The first one is reasonable because of the low level of energy, it would not survive, if it flees. Thus, the desire to flee is repressed to avoid frustration and to be able to satisfy the desire to eat for surviving. The second one is better, if the level of energy is low, but it is possible to

move to another energy source without dying. Therefore, the bubble flees and the desire to eat is repressed.

A possible technical realization could be a boolean variable which indicates if a recognized drive/desire is conscious or not, i.e. if it should be repressed it cannot be satisfied, because it is not known.

Repression is a defense mechanism, which is often used by human beings, because it is easy and flexible. Flexible in this case means that Repression can be utilized, if a drive/desire should be unconsciously either for a short time or for a long time like the whole life. So a drive/desire can be repressed because at that special moment it cannot be satisfied or is not appropriated at this time but later. So the same drive/desire is repressed at one moment and at another moment it is fulfilled.

Introjection, Identification

This defense mechanism is split in two main parts. The first part is the Introjection, in which the personality or behavior is absorbed to become the own. In this phase the bubble select unconsciously the other bubble, whose personality or behavior it would like to copy. Therefor the personality or the behavior has to be saved in a memory, but maybe also to be able to compare if it has learn well enough. The second part is the Identification, in which the personality or behavior becomes the one. In this phase the bubble tries as well as possible to imitate the other one. So this defense mechanism is difficult to realize because a lot of details have to be known, e.g. the bubble has to keep in mind which bubble it like to imitate and which personality or behavior. The behavior or the personality should be 'scanned' detailed and should be saved in such a form, to enable the imitation, as well.

Projection

When this mechanism is used, he/she describes unacceptable feelings or qualities to someone else. So to apply this mechanism it has to be saved informations about the individual, who is involved, and the feeling or the quality. An example for the BFG is a bubble, which is lazy or does nothing for the group, does not help or does not dance with another one because it believes that the other one is lazy or would not help it.

Reaction Formating

When someone uses this mechanism, he/she takes up the opposite feeling/emotion or behavior instead of the truth. An example for the BFG is that a bubble is friendly and help another bubble although it does not like the other. So for this mechanism it has to be defined which two behaviors or feelings/emotions are contrary to each other. If there is also another individuum as in the example above involved a possibility to save this is necessary.

As seen in the examples above, often another individual is involved in the process. Thus it is interesting to be able to save information about this individual. But depending on the mechanism sometimes more information has to be saved e.g. the behavior or the feeling. So it has to be thought about the possibility to save different informations and foremost how, because it should not take too much amount of memory. As well it is important that information like which two

feeling/emotions or behavior are contrary to each other can be for first implementations static but it should become dynamic because it depends on the society and therefore on the evolution state of a population.

3.3.3 Changes in Behavior explained by Use Cases

In Section 2.3 some possible influences of using defense mechanisms on special behavior are explained. In this Section the changes of the behavior due to applying one of the defense mechanisms presented in 3.1.2 are illustrated by some use cases, i.e. only the behavior, which emerges when using one of defense mechanisms, are explained. The use cases are splitted into actors, preconditions and sequence of actions. The preconditions differ between the use cases, but the precondition that the agent should have implemented the possibility to use defense mechanisms is for all use cases mandatory and thus not mentioned. The following description of the use case are based on [Koh08, p. 40-57].

Love

The three reasons for love are closeness, similarity and reciprocal affection and either lead to passionate love or to companionable love. Both kinds of love differ between their characteristics. While the first one needs to satisfy the sexual drive through displaying affection, the second wants to receive help and understanding.

The actors of this use case, regardless of which love, are two individuals. The preconditions, which have to be fulfilled by the agents, are that the agent has to have a sexual drive and the possibility to cathect the mental object of other agents with its energy. This means that modules, which handle drives and affects as well as the memory and the represention of other agents, are required. As well the agent should be able to show affection and to reduce the tension e.g. by kissing.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. At the beginning the affection for another agent increases, but it may take some time until the effects takes place. This is caused e.g. by being long time in the vicinity of each other or the agents have similarity in believes, motivations or goals. After accustomating to each other agents will be exposed to situations, which provoke an emotional excitement. This causes physical excitement which will be referred to the other agent. Now it is possible that the other one responds to the affection. This can be done in different ways, whereas only the ways are listed which causes the use of defense mechanisms. The first one is that the agent does not respond, because it is completely unaffected by displaying the affection and the vicinity of the other agent. The second one is that the agent as well does not respond or respond in an aggressive way, because it feels anticipate against the other agent. Both answers are not positive. Therefore it might be necessary to repress or sublimate the sexual drive referred to this agent. The decision for using defense mechanisms depends on the intensity of the feeling to the other agent and how often the agent is disappointed by another agent in a certain span of time. A reason therefor is e.g. the agent becomes depressive or aggressive and hurts other agents. The behavior may change in different ways. Probably it does not show its affection to another agent for a long time, which leads to the problem that the agent feels alone and not loved, or does not respond to an affection of another agent because of being afraid that the other one will go away, which attributes the agent with pain.

Aggression

Pain, frustration and provocation are the three most common factors releasing to aggressive behavior. Pain can be caused by injury or if someone feels an unease which is initiated by external or internal factors, whereas an example for the first one is too high temperature and for the second one hunger or thirsty. If someone expects the successful outcoming of something which is denied unexpected, then this can lead to frustration. Provocation is done by someone else who encourages him/her by verbally acts as well as by showing aggressive behavior against him/her.

As seen in these examples how aggressive behavior can be caused only one agent is needed for showing aggressive behavior. But the source for that can be one or more agents. The preconditions for aggressive behavior is that the agent needs a death drive or an equivalent for aggressive behavior. As well the agent has to have the possibility to attack another agent to discharge the drive tension.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. At the beginning one of the above mentioned triggers occur. After that the agent has two possibilities to respond to the trigger. The first one is to repay. Mostly the gained wrath hits the person who is the trigger for his/her aggressive behavior. But sometimes it is necessary to repress this drive because the other one has more power, e.g. the leader, or a stronger agent. This can be done by using defense mechanisms, e.g. Repression. But it is possible that the inner tension is too high and a discharging of the tension is essential. This can also be done by using defense mechanisms, e.g. Displacement which enables the person to discharge the drive tension by punching a cushion.

Stress

Various reasons for stress are presented in literature. Four of them are life changing events, catastrophes and traumatic events, chronic stressors and everyday problems. They can lead to a weakened immune system, a decreasing social life, emotional reactions like depression or an aggressive behavior because of feeling unwell.

In case of stress the number of actors can vary. When something in the environment causes stress only the agent expierienceing the stress is present. But it is also possible that one or more agents can provoke stress. A precondition for this use case is that the agent is able to put itself into a state of elevated readiness, which directs the focus of attention to objects which provokes stress. This can be helpful in dangerous situations but it makes the body susceptible on illness if it continues for a too long span of time.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. As explained in Section 2.3 an individual goes through three different states. At the beginning an agent experiences a situation which upsets itself emotional. At this time the first state named *alarm reaction* begins, because this situation alerts the agent. If the stressor is a traumatic event it may be necessary to use defense mechanisms. An example therefor is seeing an accident of a good friend where the person dies. The person who sees this, denies that this ever happens. This is done by the defense mechanism Denial. Another possible reason for using defense mechanisms is that since now the body needs all its energy to deal with the situation. If the stress continues the body starts to resist against the weakening effect of the stressor, which is the second state name *resistence*. In this state the body has to reduce all drives

to the minimum - to the drives which are necessary to survive. So in this state the drives not needed at that time are repress or sublimated to not know that they exist or to be able to use the energy for dealing with the stress. If the stress continues too long or too intensive it is possible that defense mechanisms like Conversion minimizes or deactivates bodily abilities like seeing or hearing to show that the body has no more energy to deal with the conflict.

Reciprocity

When a person gets help or a gift from someone he/she feels himself/herself obligated to repay the favor. This can be done by making the other also a present or by helping if the other is in an emergency.

In the simplest case only two agents are involved as actors, whereas the first initiates the process by setting an action, which is a favor for the other. The second one tries to pay the favor back. A precondition for reciprocity is that the agent has a Superego, whose demands tell the agent that a given favor have to be return. But also important is that it is possible for the agent to remember who did the favor to be able to repay it to the right agent as well as to know who is trustworthy.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. First an agent does a favor for another one. This can be a present or helping the other getting food or defending. After that the other agent has two possibilities to answer. The first one is no reaction on the favor. This can have different reasons but might not be a good solution because the agent which did the favor would not help in emergency again. The second reaction is responding by returning the favor. This reaction can have various causes. For example a defense mechanism evokes this reaction because of being in love with the other agent, but knowing that the other agent does only a favor if it brings benefit. Another reason of not answering the favor is that the other agent does not like this agent. But if it is necessary to keep the agent as friend for worse times, defense mechanisms like Rationalization can be useful to answer the favor although the agent would not like to do so.

Hierarchy

To see if a hierarchy is presented in a group, a small amount of a particular type of resource exist in an environment. This resource suffices only for a small number, i.e. only some of the group members benefit from it, while others achieve only a small amount or nothing. If no clear hierarchy exists in the group, members will start to fight.

A hierarchy can only emerge when agents have formed a group. As well an influence on the distribution of a resource within a group is the presence of a leader. The actors of this use case are at least two agents otherwise a hierarchy is not possible. A precondition is that the agent has to have a Ego-ideal which is filled with a superior individual. As well the demands and restrictions of the agent concerning the respect of this superior individual have to be present.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. First the agents have to form a group. This group can be predefined or it can be established over a time. Afterwards a resource should become rarely, i.e. a shortage of a resource. To be able to decide who gets the resource, a pecking order is defined in the group. This can happen in different ways like trying to intimidate each other by aggressive behavior but

not attacking. The one who gives up in first is lower in the hierarchy and receives less or nothing of the resource. After that the resource is distributed in favor of the higher group members, who will longer survive if there is a shortage of energy than the inferior group members. If an agent loses the fight for a resource it may not be easy to accept this, particularly if this happens very often. A reason therefor can be that the agent is smaller or weaker than the other. To be able to accept this defense mechanisms are used to repress the drive, which causes the desire to possess something of the resource. As well an aggressive behavior may occur if an agent often lose the fight. To be able to discharge the tension without being aggressive defense mechanisms can be utilized, e.g. Displacement for discharging the rage on an agent, which does not be a group member or on an object like a cushion. Sometimes the fight for a resource means the decision between living and dying. In such a decision it may be necessary to repress the fear and fight until the agent dies or win. In this situation defense mechanisms can be used to repress the fear or to turn the fear into aggression, which can be seen as Reversal - the one who was beaten beats someone else.

Ability to Work in a Team

Following two factors determine if an agent is able to work in a team:

1. The first one is the participation in a joint effort to reach a goal without pursuing the own one.
2. The second one is that the agent has to be able to share the reward with the group if they had helped to reach the goal. This includes that if an agent did not help, it would not gain something of the reward as well as it may happen that it is expelled from the group.

The actors in this use case are a group which means that at least two agents interact with each other. The preconditions are that the agent is member of a group as well as the agent has to have a Superego to know the demands and restrictions, which it has to obey. However, it is also possible that the agent does not follow the decision of its group but its own. This decision can be influenced by its drives as well as it can be a rational decision.

In the described type of scenario, defense mechanisms are typically applied. In the following one possible application is shown. First the goal has to be found by the group which can only be reached together. Afterwards the group works together to attain this goal which means the agents have to decide for themselves if they want to reach this goal or if they want to follow their own aim. For answering itself this question various factors have to be taken in account, e.g. is it possible to be expelled from the group if the agent does not help. At that time it is sometimes necessary to use a defense mechanism because of repressing the own drive to be able to give the goal of the group higher prioritization. As well it might be possible that another agent which will also work in the team does not be trustworthy or does not return love. So it is useful to forget this or to reversal the aggressive behavior in energy to reach the goal. This is also be done by defense mechanisms. The last action of this use case is that the reward of the effort is shared in the group. It might be possible that an agent would like to have more than it is entitled and start to fight. If an agent has too much fear but does not want to be seen as a coward it uses e.g. Rationalization which means that it offers another reason instead of the truth for fleeing. The reason might be that the other is too strong.

3.3.4 The Four Defense Mechanisms

The defense mechanisms Displacement, Sublimation, Intellectualization and Conversion are chosen for implementation in the BFG. In this Section a model for each of the four defense mechanisms is explained. These models are used to create a flow chart, which is the archetype for implementing the four defense mechanisms that is described in the next chapter.

Sublimation

As shown in Figure 3.13, Sublimation is used by searching for a solution to discharge the drive tension in a way, which is accepted in the society. With this solution the 'old' one, which is not tolerated in the society, is sublimated, to satisfy the drive. The result is the discharging of the drive tension and not a drive deferment.

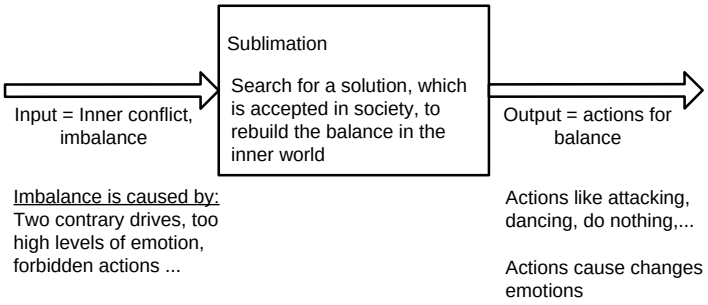


Figure 3.13: A General Model of Sublimation

Intellectualization

As shown in Figure 3.14, Intellectualization is used by searching for a solution for a problem. The solution for the problem has a logical reason as background, to discharge the drive tension. An example can be the escape because of having fear concerning the enemy. But if someone asks for the reason of the action, the person says that there had been too many enemies in visibility range. Another reason can be that the enemy was too strong to be able to win. This defense mechanism allows the discharging of the drive tension too, like Sublimation.

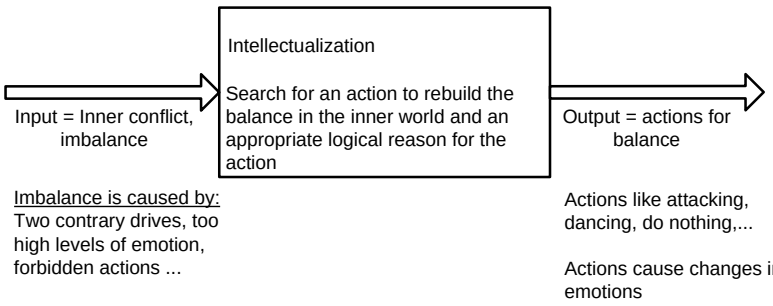


Figure 3.14: A General Model of Intellectualization

Displacement

As shown in Figure 3.15, Displacement is used by searching for an object, which has nothing to do with the primary situation, to be able to discharge the drive tension. An object can be an item as well as another human, for example someone can shout with someone else or punch a cushion for reducing the anger about his/her work. As seen in this example Displacement is e.g. utilized for reducing emotions like anger. That is the reason why also the change for them has to be considered in the model of this mechanism. Displacement enables human to discharge the drive tension as well as Intellectualization and Sublimation.

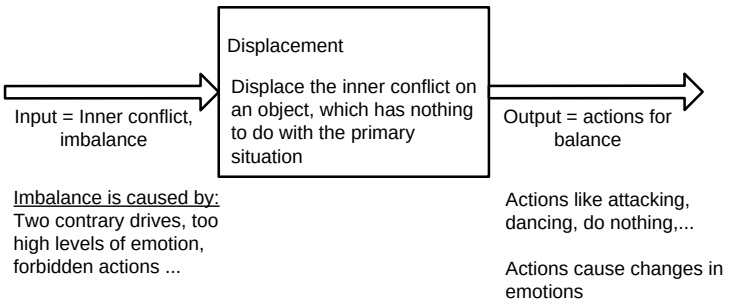


Figure 3.15: A General Model of Displacement

Conversion

As shown in Figure 3.16, Conversion is used by transforming the inner conflict into a somatic symptom. An example for a somatic symptom can be not seeing or not hearing. The first one may occur because somebody wants to see forbidden situations as an example a voyeuristic person. In this model it is shown that the output is only the somatic symptom and no discharging of the drive tension happens. This mechanism influences a bodily function like seeing, hearing or moving arms and if it is used for autonomous agents, it should be possible to reverse this like in the human world.

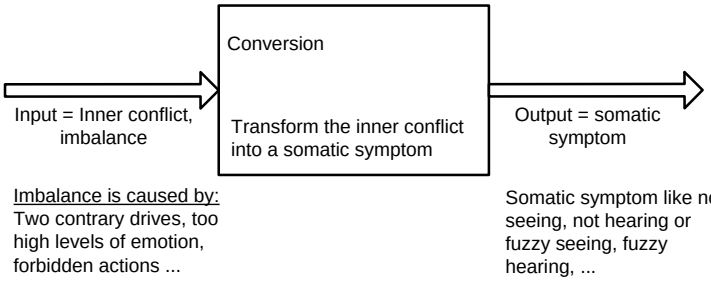


Figure 3.16: A General Model of Conversion

Summarized for chapter 3 the theory of psychoanalysis has been explained in detail. First the second topographical model of S. Freud was described, on which the theory of the defense mechanisms is based. Afterwards various defense mechanisms are presented and examples of the human world are analyzed. Then a general model is developed, which is transformed into a technical model for implementation. Furthermore selected defense mechanisms are explained and analyzed.

As well some common properties of the defense mechanisms and important informations for technical implementations are described. Finally a general model for the four defense mechanisms Sublimation, Intellectualization, Displacement and Conversion is presented. In the next chapter the technical realization is described. First some basics like the used "Structured Scenario Recognition System" or the cycle in the function doThinking, in which the defense mechanisms are implemented, are explained. Afterwards the implementation of the four defense mechanisms Sublimation, Intellectualization, Displacement and Conversion and the prioritization is described in detail. Finally, the test benches and the simulator are shown and the general test conditions are presented.

4 Technical Realization

For the technical realization the four defense mechanisms Sublimation, Intellectualization, Displacement and Conversion are chosen because of being applicable for implementation in the Bubble Family Game (BFG). First basics for the technical realization are explained. There the general theory about the structured scenario recognition system is described, which is defined by using an XML structure, as well as the function *doThinking*, where e.g. the recognition of scenarios and the control for using the different defense mechanisms is placed. Then a general cycle for the four defense mechanisms is presented. Afterwards an example of each defense mechanism, which is implemented in the BFG, and the implementation of the example is explained in detail. As well a description of the prioritization of the defense mechanisms is done. Furthermore, the editor for the XML files of the defense mechanisms is explained. Finally, the test benches and the simulator as well as the different test conditions are presented.

4.1 Basics for the Technical Realization

In this Section some basics for the technical realization are presented. The implementation of the defense mechanisms is done by using XML files for structure definition and java classes, including the corresponding logic. For the XML-files the style-sheets of AbstractImages, Scenarios and Desires, which already have been defined in the project ARS-PA, are also used for the defense mechanisms. In the following the design of the structured scenario recognition system is explained in detail. To see, where the defense mechanisms are located in the Java code, the function doThinking is explained, which is responsible for the recognition of images, scenarios and desires as well as other task.

Structured Scenario Recognition System

The scenario recognition model, which is implemented in the BFG, is based on predefined perception patterns, called image templates, and predefined pattern of possible scenarios. The image templates combine different sensor outputs and gives them a semantic meaning, but they are a set of perceived data within a single moment. Recognized images are utilized as transition conditions between states of a scenario recognition process - scenarios are implemented as state charts. As contrast to images, scenarios are a perceived sequence in time. Important for this model is that the image templates and the scenarios can be predefined. Desires are derived from scenarios, but the difference to scenarios is that transition conditions can also be performed actions [LBP+07].

During the development of the new model of the project ARS Abstract Images are renamed by Mental Images.

The components of an image template are rules, which defines the perception of specific types of symbols. Symbols can be very simple like the temperature in a room or contain complex information such as 'a meeting is taking place in the conference room' and are following the work of [Pra06]. To define the concept of image templates first in a generic way, the name and meaning of symbols shall be abstracted and are named in the further text S_1 to S_m (shown in Formula 4.1), but also some examples are explained later.

As seen in Formula 4.2, the module of such a system produces a constant stream of symbols, so that in every calculation step the set P (set of perceived symbols) contains a subset of symbols of the set S (set of all possible perceived symbols). The image template (shown in Formula 4.3) is a subset of the set of perceived symbols P, but the symbols are weighted. To be able to compare all image templates with the perceived and weighted set of symbols, a tree structure was created, in which the content of one image template is defined.

$$S = \{S_1, S_2, \ldots, S_m\} \;\; possible\, perceived\, symbols \tag{4.1}$$

$$P = \{S_4, S_6, \ldots, S_l\} \subseteq S \;\; perceived\, symbols \tag{4.2}$$

$$IT_1 = \{w_4 * S_4 \; AND \; w_6 * S_6\} \subseteq P \;\; Image\, template \tag{4.3}$$

The tree has three basic elements (seen in the Figure 4.1) - *image element* (iE), *image node* (iN) and *image leaf* (iL):

NAME	SYMBOL	BASE
iE		-
iN		iE
iL		iE

Figure 4.1: Basic Elements in Image Template Definition [LBP+07]

The basic class of all elements is the *image element*, which holds the following data. The element has a name, optional a description of the specific use and the information if this element is mandatory or optional. An *image node* is equipped with informations about its child nodes, which can be further *image nodes* as well as *image leafs*. It holds information about the composition (AND or OR) between its sub nodes and a negation flag that represents a boolean negation of all sub elements. In the *image leaf* it is defined which type of symbol shall be part of the image template detection and a logical operator (==,!=,<,<=,>,>=). With this logical operator the value of a symbol is compared to the value defined in this *image leaf*.

An example of a generic image template definition is shown in Figure 4.2 and described bottom to top. The value of symbol S_2 has to be equal or higher than 'HOT' (as a symbolic output for e.g. a temperature sensor) OR the value of symbol S_3 has to be equal or lower than 'DAMP' (as a symbolic output for e.g. a humidity sensor). This is the first half of the image template and is matched if the perceived data is within the range of values. The second half is matched, if the value of the symbol S_1 is equal or higher than 'NORMAL'. Because of the logical operator AND in the base node, the image template only matched fully, when both conditions of the sub nodes *LEAF 1 incl. S1* $\rightarrow$ *VALUE(*$S_1 \geq$*'NORMAL')* and *SUB NODE, "OR"* are fulfilled.

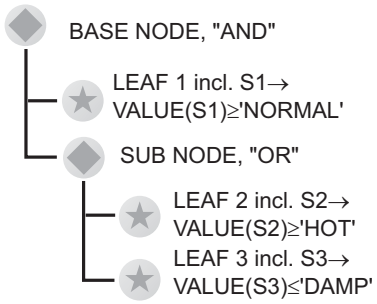


Figure 4.2: Example for a Generic Image Template Definition Accomplished in a Tree Structure [LBP+07]

The quality of the match is a value, which can be between zero (no match) and one (100 percent matched). To calculate this quality value, a simple algorithm has been implemented, which compares the predefined image template with the actual perceived symbols. First the total weight of the tree is calculated and it does not matter if there is a match or not. Therefor all leafs are counted, but the leafs within a node containing a logical operator OR are seen as one leaf and the leafs within a node containing a logical operator AND are all counted. This is done through all branches of the tree. Then all leafs, which have a valid condition match, are count (= number of matching). The quality of match is equal to the number of matching divided by the total weight of the tree.

MATCH	LEAF 1	LEAF 2	LEAF 3
1.0	X	X	-
1.0	X	-	X
0.5	X	-	-
0.5	-	X	X
0.5	-	-	X

Figure 4.3: Possibilities of Matches in the Example Image Template [LBP+07]

In Figure 4.3 the quality of matching is seen, whereas the tree in Figure 4.2 is used for calculation. The total weight of the tree in the example above is two - one for leaf 1 and one for leaf 2 or 3 or for both. If only leaf 1 is recognized, the quality of match is 0.5 (seen in row three in Figure 4.3, which is one match divided by two (= total weight of the tree). For the realization of such image templates an XML structured file is used.

Abstract Images can be situations, comparison of levels of emotions with a threshold etc. All important Images for the different defense mechanisms are saved in one file. If an Image should evoke an action an additional XML file has to be designed, in which the action is assigned to the appropriated Image-ID. An example of an XML code for an Image is shown in Figure 4.4, which shows a comparison of the level of the drive play with the threshold „High“. If this condition is fulfilled the action - shown in Figure 4.5 - is loaded in the Action Container. When the priority is high enough, the action is carried out. As shown in Figure 4.5, there is a parameter called *ID*, which identifies the action - in the example 42 means that the bubble asks its teammate if someone wants to dance with it. The parameter *Source* defines if the action is an interaction with the

environment or another bubble (*Source* = outer, which means the environment) or a distribution of hormones/antihormone (*Source* = inner, which means the body), which influences the level of emotions. If the hormone is injected, the level of the emotion rises and if the antihormone is injected, the level of the emotion falls. There is also a parameter called *Priority*, which is important for the prioritization of the outer actions. For the inner actions the value of the *Priority* is used for calculation of the level change of the emotions. This value is only between 0 (equal to 0%) and 1 (equal to 100%), whereas the value 1 for outer actions means the highest priority and 0 the lowest priority.

```
- <AbstractImage ID="3" name="Desire to dance" fullMatchRequired="true">
    <Description>Desire to dance</Description>
  - <TreeRoot optional="mandatory" booleanOperator="AND" negated="false">
    - <leafDrive compare=">=" level="High">
        <drive>play</drive>
      </leafDrive>
    </TreeRoot>
  </AbstractImage>
```

Figure 4.4: Example for an Image in XML

```
- <AbstractImageAction AbstractImageID="3">
    <Description>Ask for dance</Description>
  - <ActionPlan>
    - <Action ID="42" Priority="1" Source="outer">
        <Name>Ask for dance</Name>
      </Action>
    </ActionPlan>
  </AbstractImageAction>
</AbstractImageActions>
```

Figure 4.5: Example for an Action in XML

Every image and the resulting match does not contain the variable time, they deal only with a single moment. For the concept of scenarios, the value of time is added, so the perceived data of the past is now also included. So a scenario is a sequence of several recognized images in time, which are perceived. It has been decided to use the concept of state charts for the representation of such scenarios. In the scenario recognition process may occur several ways from the beginning to the end, e.g. there may be more than one event that must be perceived by the image template recognition. Also different paths are possible for full recognition of the same scenario, e.g. as shown in Figure 4.7 there are following two paths possible:

1. First path:
   The recognition of first IT 1, afterwards IT 2 and finally IT 5.
2. Second path:
   The recognition of first IT 1, afterwards IT 3, then IT 4 and finally IT 5.

The scenario recognition process may also be aborted, e.g. by a time out or another event. So it is necessary to be able to make global abort conditions possible. In Figure 4.6 the basic elements of a scenario definition and their purposes are shown, whereas the start and the end states only have a virtual meaning [LBP+07].

All scenario recognition processes start at the begin state and wait for the first transition condition. When the end state is reached, the lifetime of the scenario recognition process ends and no

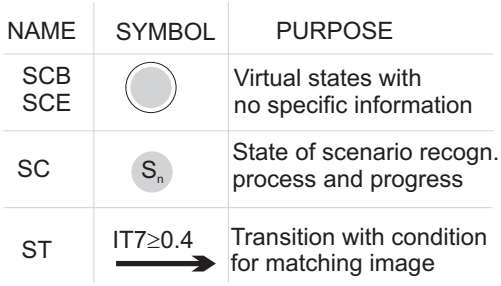


Figure 4.6: Basic Elements of Scenario Definition [LBP+07]

further information is included. All states except the end state holds a list of transitions, which specify the condition to switch to the next state and has to be met by the match of a selected image template. As seen in the Figure 4.6 the condition for switching to the next state is IT 7 >= 0.4, i.e. the image template 7 has to match for 40 percent or higher.

In Figure 4.7 a generic scenario template definition can be seen. To switch from the begin state to the state S_1, the image template IT 2 has to be perceived with a match at least 60 percent. If this happens, the current state is set to the state S_1, which has the following two different transitions:

1. The image template IT 2 has to be perceived with a match of hundred percent, then the next current state will be S_2.
2. The image template IT 3 has to be perceived with a match of at least 80 percent, then the current state will be S_3. If afterwards the image template IT 4 is perceived with at least 70 percent, the path is closed and the recognition process gets also in the state S_2.

To fulfill the final condition for complete scenario recognition, the perception of the image template IT 5 with a match of at least 50 percent has to be done.

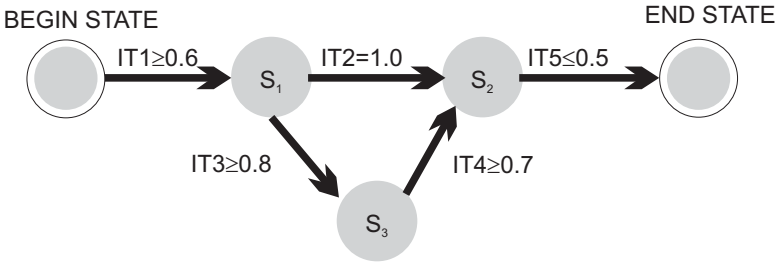


Figure 4.7: Generic Scenario Template Definition [LBP+07]

For the realization of such scenario templates an XML structured file is used. Both, image templates and scenario templates, represents the knowledge of the system. Images and scenarios, which are not predefined in this database, will not be recognized.

The XML file for a scenario is used for the recognition of the exceeding of the levels of emotions or drives over a threshold, e.g. for Sublimation that the level of the emotion rage is equal or over high. If the condition is fulfilled the XML file for the core function of the defense mechanisms, which has the desire XML file as archetype, is initiated. Only the setting of the fuzziness of the

sensors and the prioritization of the defense mechanisms is programmed in Java. In Figure 4.8 an example for an XML file of a state in a scenario is shown. The basic structure of this XML file is a state chart. The state has an ID, which is also the key name. The condition is an Abstract Image, which has to be recognized, and as target state the ID of the next state is presented. As seen in this example a state can have more than one target state.

```
- <State ID="1">
    <Name>Drive levels too high</Name>
    <ProbabilitySuccess>0.001</ProbabilitySuccess>
    <StepsToTimeout>1</StepsToTimeout>
    <DecreaseDelta>0.03</DecreaseDelta>
    <MaxTimeouts>1</MaxTimeouts>
  - <Transition>
      <TargetState>2</TargetState>
    - <Condition>
        <AbstractImageId>17</AbstractImageId>
        <MinLevel>High</MinLevel>
      </Condition>
    </Transition>
  - <Transition>
      <TargetState>2</TargetState>
    - <Condition>
        <AbstractImageId>18</AbstractImageId>
        <MinLevel>High</MinLevel>
      </Condition>
    - <AbortCondition>
        <AbstractImageId>28</AbstractImageId>
        <MinLevel>Medium</MinLevel>
      </AbortCondition>
    </Transition>
  </State>
```

Figure 4.8: Example for a State of Scenario in XML

As explained in Section 2.2, it is possible to abort a scenario/a desire or that the current state falls back to a state before. The first one is realized as an abortion condition in the transition, which is a recognized image. The second one is implemented by setting the variables *StepsToTimeout* and *MaxTimeouts*, whereas *StepsToTimeout* is the number of failed transition changes to initiate a time out and *MaxTimeouts* is the maximal number of timeouts, which causes a state change to the state before (falling back to the state before). Both is seen in the example in Figure 4.8.

The Function 'doThinking'

The function *doThinking* is located in the class *clsBrain* and is the main function of the decision unit implementation. It is e.g. responsible for the recognition of images, scenarios and desire or for planning the next actions, which have to be set. In the following the cycle in the function *doThinking* is described in detail. First the recognized AbstractImages are load in the container *oCompareResults*. As explained in Section 4.1, images can include actions. These actions are loaded in the containers *oOuterActionList* and *oInnerActionList*. That means that inner and outer actions are splitted into two different containers, because all inner actions are executed but the outer actions are sorted by the priority before the action with the highest priority is executed. Afterwards the function *triggerScenarioRecognition(oPrioritization, oChangedScenarios, oRecognizedResult, oInnerActionList, moEmotions)* is called. There the scenarios, which are initiated by the recognized images, are loaded in the container *oRecognizedResult*, if these images

are the first condition. If the first condition had already been recognized, the scenarios are in the container and only a state change is carried out. If a full recognized scenario initiates a desire, it is loaded in the container *oChangedScenarios*. Before this happens, the prioritization of the defense mechanisms is done.

This is the first part that is necessary for the implementation of defense mechanisms. The logic is situated after the scenario recognition, because the scenarios for the defense mechanisms are now recognized. But if two are initialized, only one of them should be initiated a defense mechanism as well as only the actions of this one should be loaded in the containers. After the prioritization the actions of the loaded desires and defense mechanism are added to the global action list, whereas the inner actions are loaded in the container *oInnerActionList* and the outer actions in the container *oOuterActionList*. At this place the second part that is necessary for implementation of defense mechanisms - the fuzziness of the senors. Here it is checked first if the needed scenario is recognized and if this is done, a new inner action is created and loaded in the action container *oInnerActionList* to set the sensor value fuzzy. Finally it is checked if the Superego forbids an action, which causes a conflict and an innerconflict action is generated.

4.2 The Implementation of the Defense Mechanisms

In this Section, the four defense mechanisms Sublimation, Intelletualization, Displacement and Conversion are explained in detail. First a general cycle of each is described to show how the mechanism works. Finally an example and the implementation in the BFG is presented.

4.2.1 Sublimation

The defense mechanism Sublimation is used to sublimate unacceptable feelings/emotions or behavior in a way that is tolerated by the society. An example is the reduction of rage by an action, which is acceptable e.g. boxing.

In Figure 4.9 a general flowchart shows how Sublimation can be realized. First an inner conflict occurs, which cannot be solved. Then the Ego searches for a feeling/emotion or behavior, to sublimate the unacceptabled one with another one, which is tolerated by the society. Because of finding more than one, the Ego decides, which one is at that moment the best solution. In the BFG an example of Sublimation is the sublimation of rage, which can cause an attack of a teammate, with dancing. So the bubble tends to dance instead of attacking another bubble.

The implementation is done with two XML structured files, whereas the first one has as style sheet the scenario and the second one has as style sheet the desire. The second one is necessary, because in the scenario no action can be set and the first one is required because a desire is initiated by a recognized scenario. In the scenario file the detection if the level of the emotion rage is over a threshold is done. If this condition is achieved, the scenario is full recognized. As a result of that it initiates the second one, in which the action dancing is set. While the bubble dances the level of the emotion rage falls, which is caused by an inner action. If the level is under a defined threshold the bubble stops to dance. For the defense mechanism Sublimation an abortion condition is implemented. This is done because the bubble would dance until the level is under a threshold, i.e. it could not set another action, which may be more appropriated in some situations. For the implementation in the BFG the condition is fulfilled, when the level of the

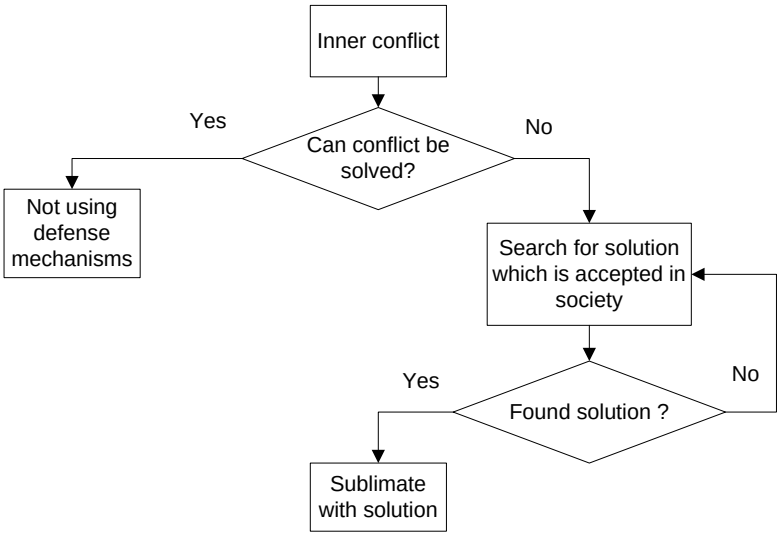


Figure 4.9: A General Flowchart of Sublimation

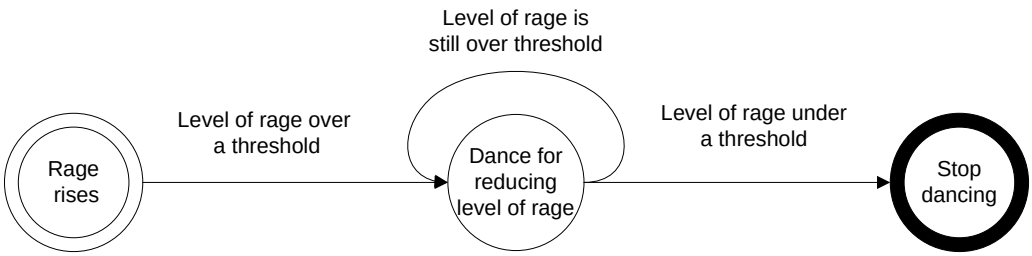


Figure 4.10: An Example of Sublimation in the BFG

emotion fear is over the threshold High, i.e. if the bubble is terrified the mechanism is aborted. This establishes the change that if many enemies are seen by the bubble and the level of rage is at that time not important enough, because an attack is more dangerous as a high level of rage. So by abortion of dancing the priority is switched to another more important action, e.g. to flee or if the bubble is attacked, it is able to defend.

4.2.2 Intellectualization

The defense mechanism Intellectualization is used, when the action is based on a rational/logical reason for the appeared problem, which is seen as the solution for the inner conflict.

In Figure 4.11 a general flowchart shows how Intellectualization can be realized. First an inner conflict occurs, which cannot be solved directly. So a logical reason is searched for the next action, which is set when this is found. The reason is important for the person himself/herself, but also to have an answer if someone of the society asks, why he/she acts like that.

The implementation is done with two XML structured files. The reason therefor is the same as explained in Section 4.2.1. The first one is for the recognition that a teammate needs help for defense and the full recognized scenario initiates the desire defined in a second XML file. In the second file the level of the emotion fear is compared with the threshold and if the level is over

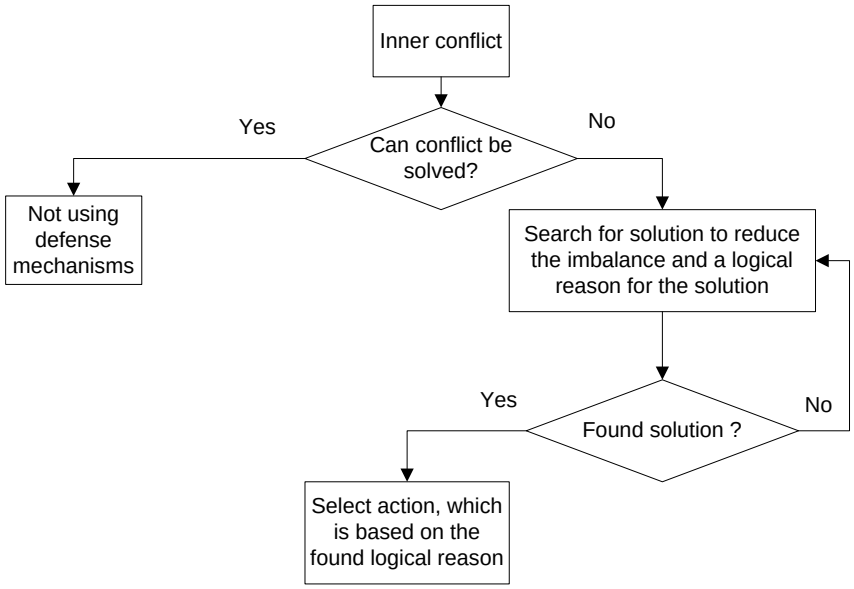


Figure 4.11: A General Flowchart of Intellectualization

a threshold the bubble flees instead of helping its teammate. The logical reason therefor is that there are too many enemies or the other bubble/s are too strong and to be able to survive it is better to flee instead of fighting. For this defense mechanism no abortion condition is implemented because it is only set one action, which does not take that long and does not disable other actions for a long time.

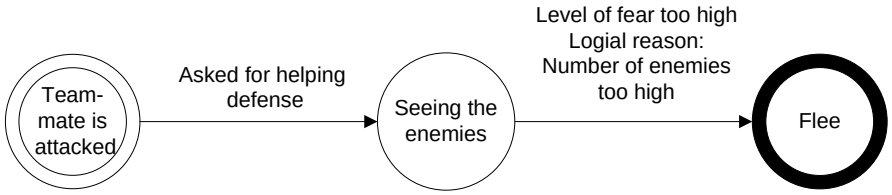


Figure 4.12: An Example of Intellectualization in the BFG

4.2.3 Displacement

The defense mechanism Displacement is used to discharge feelings like rage on objects, which have nothing to do with the primary situation.

In Figure 4.13 a general flowchart shows how Displacement can be realized. First an inner conflict appears and cannot be solved, this means that a discharge of drive is not directly possible. For example emotion disappointment rises and hope falls because of not getting help. If the levels of drive tension and repression are too high, an action is set. An action can be active, e.g. attacking people, or passive, for example someone does not help if someone else begs him/her for help. But important is that the tension of the drive, which is charged by an inner conflict, will be discharged on an object, which has no relationship to the primary situation. For example a person has problems at work, e.g. with the boss, but is not able to deal with it and cannot speak

with his/her boss. So an inner conflict is caused and the person becomes frustrated. If the level of frustration is over a threshold, a discharging is done by shouting another person at home or somewhere else.

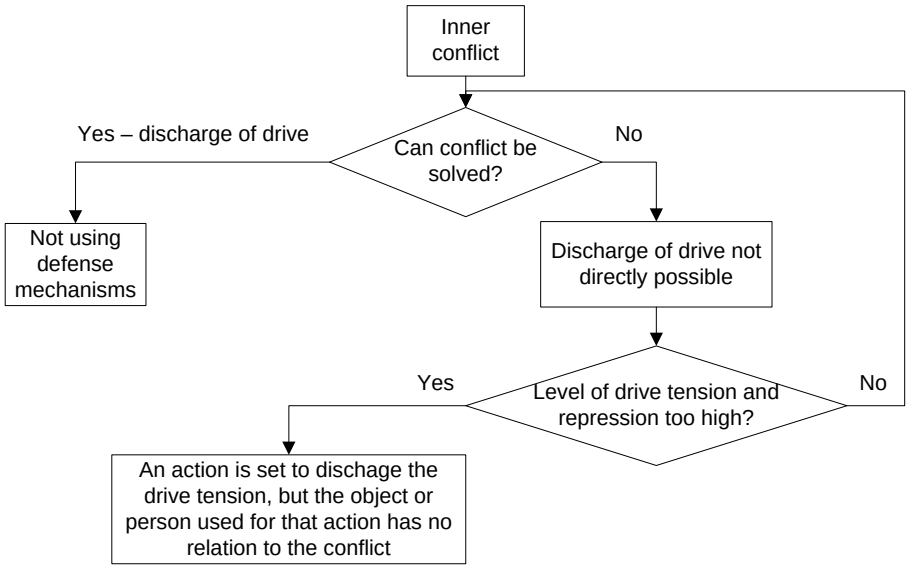


Figure 4.13: A General Flowchart of Displacement

A possible implementation of the Displacement in the BFG can be seen in Figure 4.14. The first part of the Figure is the event chain which should happen before the defense mechanism Displacement is used. There exists two possibilities to activate the defense mechanism. The first is on the left side, here a bubble is hungry and wants to eat. But the nearest energy source can only be 'cracked' with another bubble, i.e. the bubble needs help. The second one is on the right side, here a bubble would like to dance with another bubble. In both situations it asks other bubbles of its team. If the bubble receives no answer it cannot satisfy the desire to eat/the desire to dance. For this reason the emotions rage and disappointment rise. If a threshold is exceeded the bubble uses Displacement, which is shown in the second part of the Figure. Thus, another bubble of the other team or of the own team, which was not be able to listen the question, will be attacked, if it was seen by the aggressive bubble. Another possibility to discharge the tension of the drive is, that the aggressive bubble does not help another bubble, which is asking for help.

The implementation of Displacement in the BFG is done with two XML structured files. The reason therefor is the same as explained in Section 4.2.1. The first one is for the checking that a desire of eat or dance exists. In the second one the rest of the above explained scenario is implemented, which starts with the action 'asking for help' and ends with the discharging of the tension by setting an inner action when the action 'do nothing' or 'attacking' is set.

4.2.4 Conversion

The defense mechanism Conversion is used to make the inner conflict observable in form of a somatic symptom, but the person does not know the relation between the symptom and the conflict. Conversion is a really dramatic mechanism and to help people, who utilize this, many

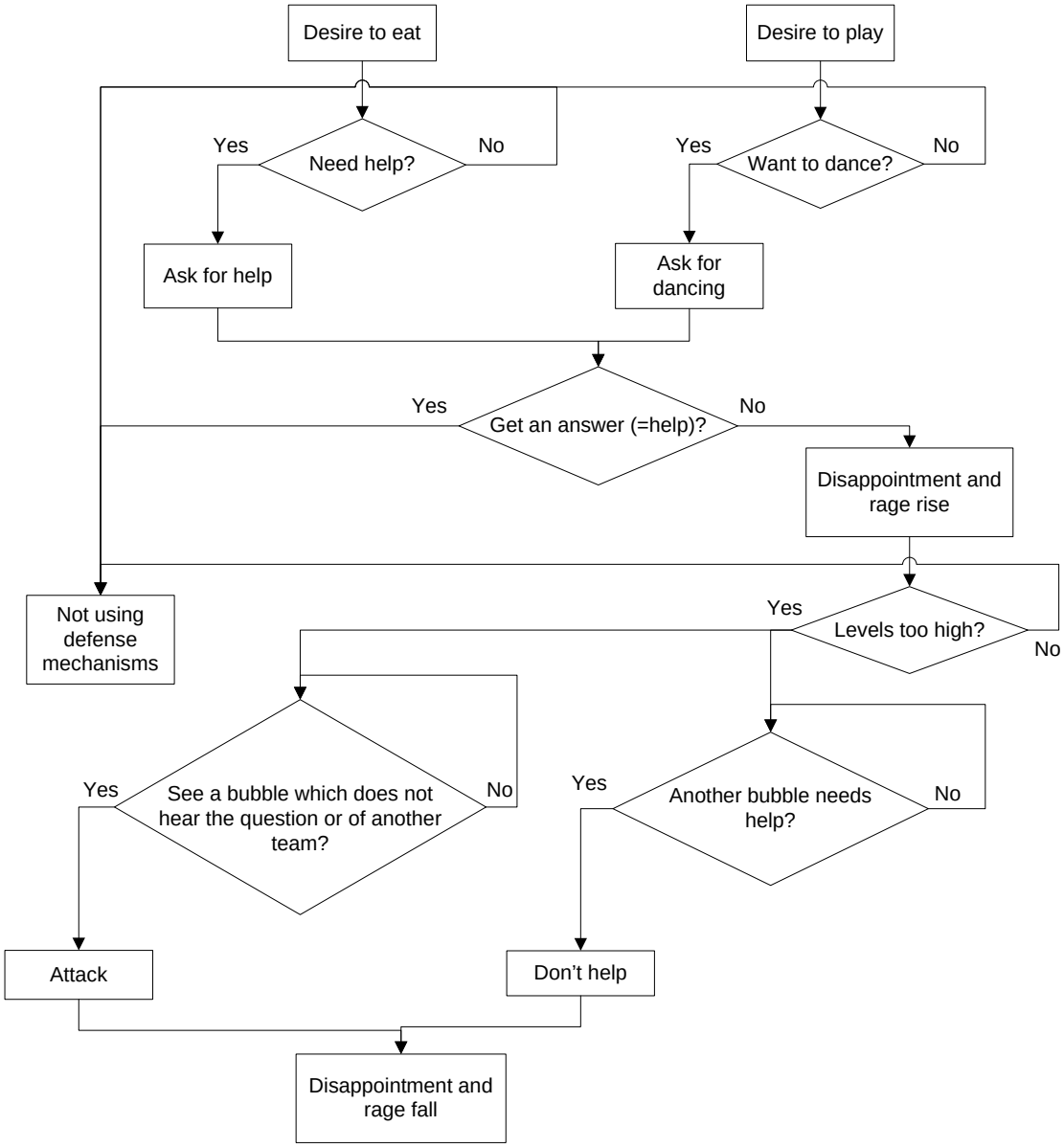


Figure 4.14: A Flowchart of an Example of Displacement in the BFG

time of therapy has to be spent in. The abolishment of this mechanism is also considered in the model. An example is if a person cannot see because he/she would like to see something, which is unacceptable in society or in his/her group, e.g. a voyeur.

The flow chart of Conversion is shown in Figure 4.16. At the beginning an inner conflict occurs, which cannot be solved respectively the discharge of the drive tension is not directly possible. If the level of the established tension is too high, the body reacts in form of a somatic symptom like not seeing or hearing respectively seeing fuzzy.

A possible implementation of Conversion is shown in Figure 4.17. Conversion is used, if more than two drive levels have arisen over a given threshold. This threshold is influenced by emotions

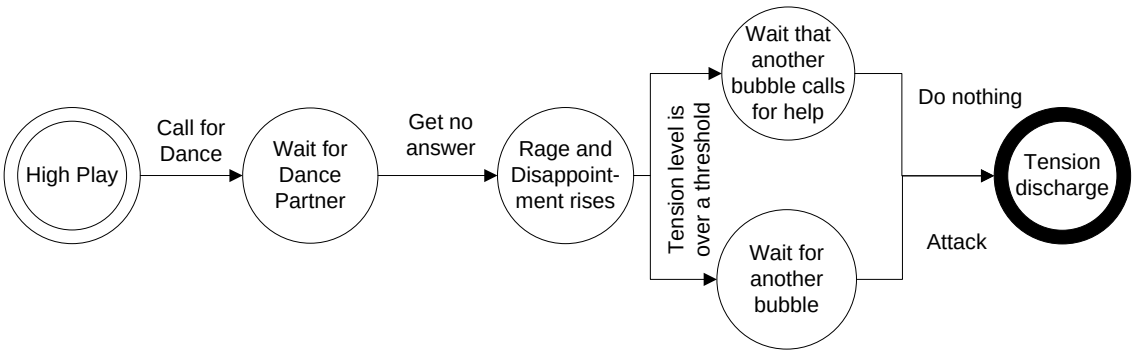


Figure 4.15: An Example of Displacement in the BFG

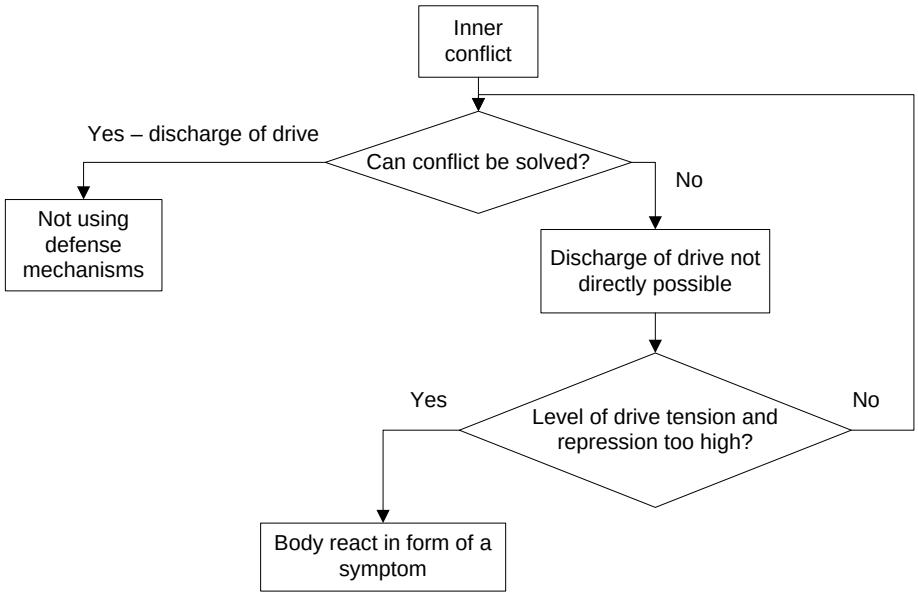


Figure 4.16: A General Flowchart of Conversion

like hope or disappointment. For example a bubble, whose level of hope is very high does not give up so easily and tries longer to find someone for helping to 'crack' the energy source for eating as a bubble, whose level is very low or whose level of disappointment is very high. If the levels of the desire to eat and the desire to play exceed over a given threshold, the bubble is not able to see exactly, but fuzzy. The fuzziness of the sensor data is also dependent of the levels of emotions, e.g. rage, disappointment, hope and joy. To implement, that this defense mechanism is not easily to reverse, different ways are possible. A feasibility is to make the reversion dependent of the help of other bubbles, i.e. a bubble needs very often help from its team members to reduce the fuzziness of the sensor. Another possibility, which is used in this realization, is that a reduction of the fuzziness of the sensor is done continuously with the time, i.e. a sloping function is responsible for the value of the fuzziness.

The implementation of Conversion - shown in Figure 4.18 - in the BFG has two parts. The first one is defined in an XML file, which is taken for the detection of levels of the desires to eat and to dance. The second one is the setting of a variable to realize the fuzziness of seeing, which

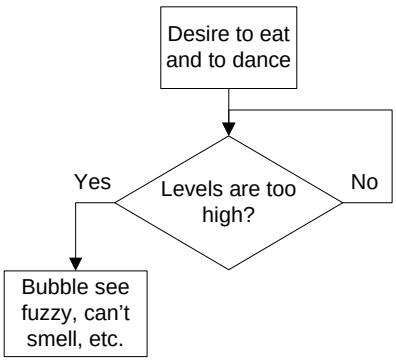


Figure 4.17: A Flowchart of an Example of Conversion in the BFG

is directly implemented in Java-Classes in the BFG - project. The function that is responsible for, starts with a value which increases during the time, i.e. it is a rising function. In the BFG the value for seeing fuzzy is between zero and one. But one, which is the maximum, means that the bubble sees all clearly (=without fuzziness) and zero means that the bubble sees nothing. Thus the implemented function f(t), seen in Formula 4.5 is a rising function, which means that the fuzziness starts at a value smaller than one and arises after a time till the value is one. If the value of the function f(t) is higher than one, it is set to one. As shown in Figure 4.18, these following four ways to initiate the defense mechanism Conversion are implemented:

1. If the level of rage is *Low*, the level of the drives lust and hunger have to be *VeryHigh*.
2. If the level of rage is *Normal*, the level of the drives lust and hunger have to be greater or equal *High*.
3. If the level of rage is *High*, the level of the drives lust and hunger have to be greater or equal *Normal*.
4. If the level of rage is *VeryHigh*, the level of the drives lust and hunger have to be greater or equal *Low*.

As shown above, the higher the level of rage is the lower is the threshold for the drives lust and hunger for activating the defense mechanism Conversion. An abortion of the defense mechanism is not necessary because it causes only an inner action and therefore the setting of an outer action is not influenced.

The following two formulas are implemented, whereas the function f(t) calculates the level of the value for seeing fuzzy:

$$z = (level_{hope} + level_{joy} - level_{disappointment})/12 - level_{rage}/7 - 0.5 \tag{4.4}$$

$$f(t) = 1 + z + (0.6 - z) * t \quad with \quad 0 \leq f(t) \leq 1 \tag{4.5}$$

As seen in the formula 4.4 for z the levels of the 'negative' emotions like rage are subtracted and the levels of the 'positive' emotions like joy are added. The reason therefor is that z should be more negative if the levels of the 'negative' emotions are higher that the levels of the 'positive' emotions, because the function f(t) has its maximum at one (100% seeing). The value z should minimized the value of f(t), because the smaller f(t) the more the bubble sees fuzzy. In the formula

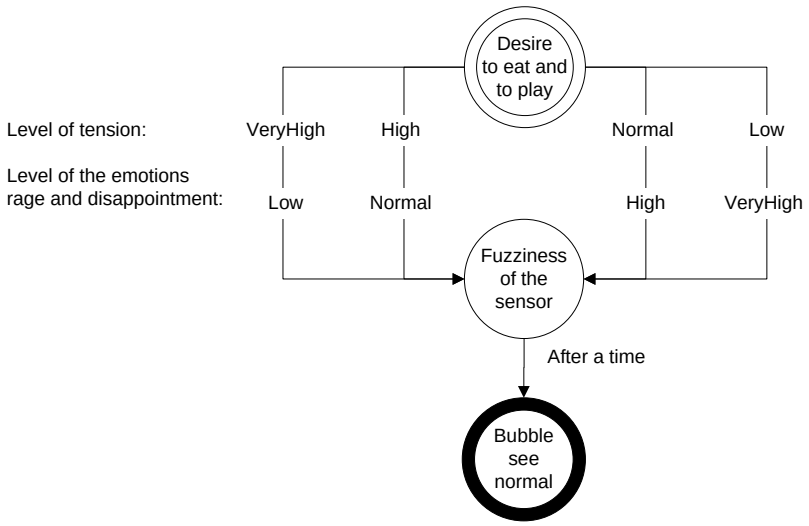


Figure 4.18: An Example of Conversion in the BFG

4.4 the subtraction of 0.5 is because of having a fuzziness if the levels of the 'positiv' emotions are higher than the level of the 'negativ' emotion and is exact 0.5 because the maximum of the first part (addition and subtraction of the levels of emotions devided by 28) is nearly 0.4, which means the minimum of z is approximately -0.1. Because the range of the fuzziness f(t), which is a number between zero and one, the level of all emotions (level of the basic emotions are between 1 and 7, the level of the complex emotions are between 1 and 4; this is explained in detail in Section 2.2) has to be normalized. This is done with the devision of 12 and 7 respectively, which is the multiplication of 3 (number of included emotions) and 4 (number of possible levels for the emotions) respectively the multiplication of 1 (number of included emotions) and 7 (number of possible levels for the emotions). The fuzziness f(t) 4.5 is the summation of 1 (maximum, bubble sees all at the right position), z and $(0.6 - z) * t$. The higher z is the smaller is f(t), which means that the bubble sees more fuzzy. The last summand is for the influence of time, so if more time is elapsed the bubble sees less fuzzy. This is the implementation of the possibility of getting help in human world. So the bubble sees better when more time is elapsed. But this is done slowlier if the levels of the 'negative' emotions are higher than the levels of 'positive' emotions. This is included with the subtraction of z in the term $(0.6 - z) * t$. The variable t (=time) has the unit minutes.

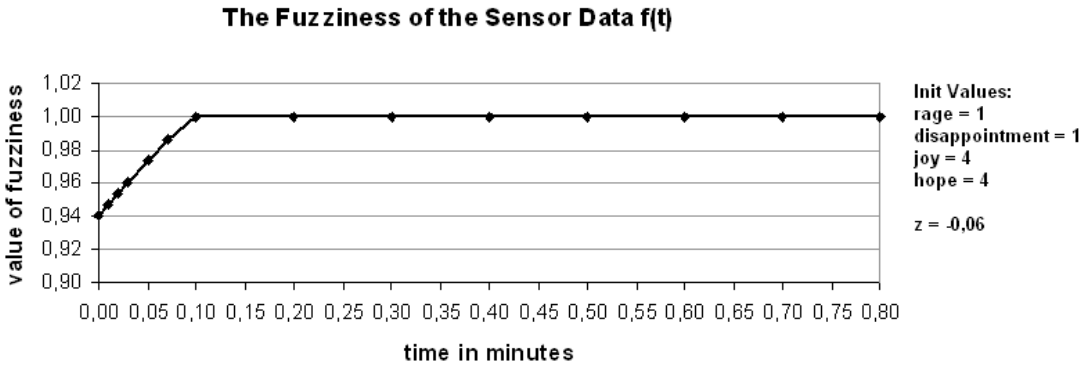


Figure 4.19: Value of Fuzziness in Dependence of Time

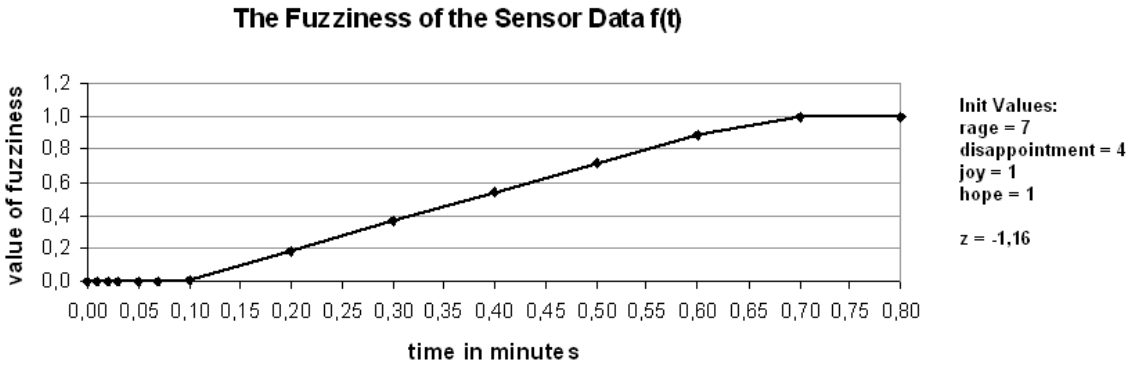


Figure 4.20: Value of Fuzziness in Dependence of Time

In the first example the level of the 'positive' emotions are maximum and the level of the 'negative' emotions are minimum. There the time for seeing normal (=100%) it takes 0.05 minutes. The trend of the function f(t) is shown in Figure 4.19. In the second example the level of the 'negative' emotions are maximum and the level of the 'positive' emotions are minimum. The time for seeing normal (=100%) takes 0.7 minutes. The trend of the function f(t) is seen in Figure 4.20. This shows that it takes more time for seeing normal when a bubble is more frustrated.

Because not all important functions are included now in the BFG, following has to be programmed in further work to use this defense mechanism, which is also shown in the right order of information flow in Figure 4.21:

1. The setting of the blur value is done with an inner action, which has to be defined in the class *enumTypeBrainInnerAction* in the Anylogic-Project-File *anyEnum*.
2. The variable, which is used for blurring, is defined in the class *clsPsychoHormones* in the Anylogic-Project-File *anyEntity*. In this class also the inner action is initiated, when it is load in the action container. Therefor the function *performAction(int pnActionId, float prEfficiency)* is responsible, whereas the variable *pnActionId* is the identification of the action and the variable *prEfficiency* contains the value of seeing fuzzy.
3. If the hormone is injected, the blur value is saved in the variable *mrQuality* of the class *clsEntityBubblePsycho*. This value is used for sending to the World Simulator, which is necessary because of the isolation between Entity Agent and World Simulator (explained in detail in Section 2.2) - so the simulator has to know that the value should be blurred.
4. The sending of the message to the World Simulator is done in the class *clsEntityBaseSensor* in the Anylogic-Project-File *anyEntity*.
5. The message is built in the package *pkgSensor*, wherefore the message has to be extended with all quality values for each sensor (There exist e.g. a sensor for StandingAboveEnergySource, which is a boolean value, or a sensor for locating other bubbles, which is a list of entries).
6. In the Anylogic-Project-File *anyPhysicsSimulator* the function *handleMsgAction* is used to read the blur value of the message and saves it there, to blur the values before sending to the Entity Agent back.
7. In the Anylogic-Project-File *anySensors* all sensor are defined and with the function *setQuality(float prQuality)*, which has to be implemented, the blur value is set with the value

achieved in the message. Some sensors like the radar-sensor is a data list and thus it is important that all entries are blurred.

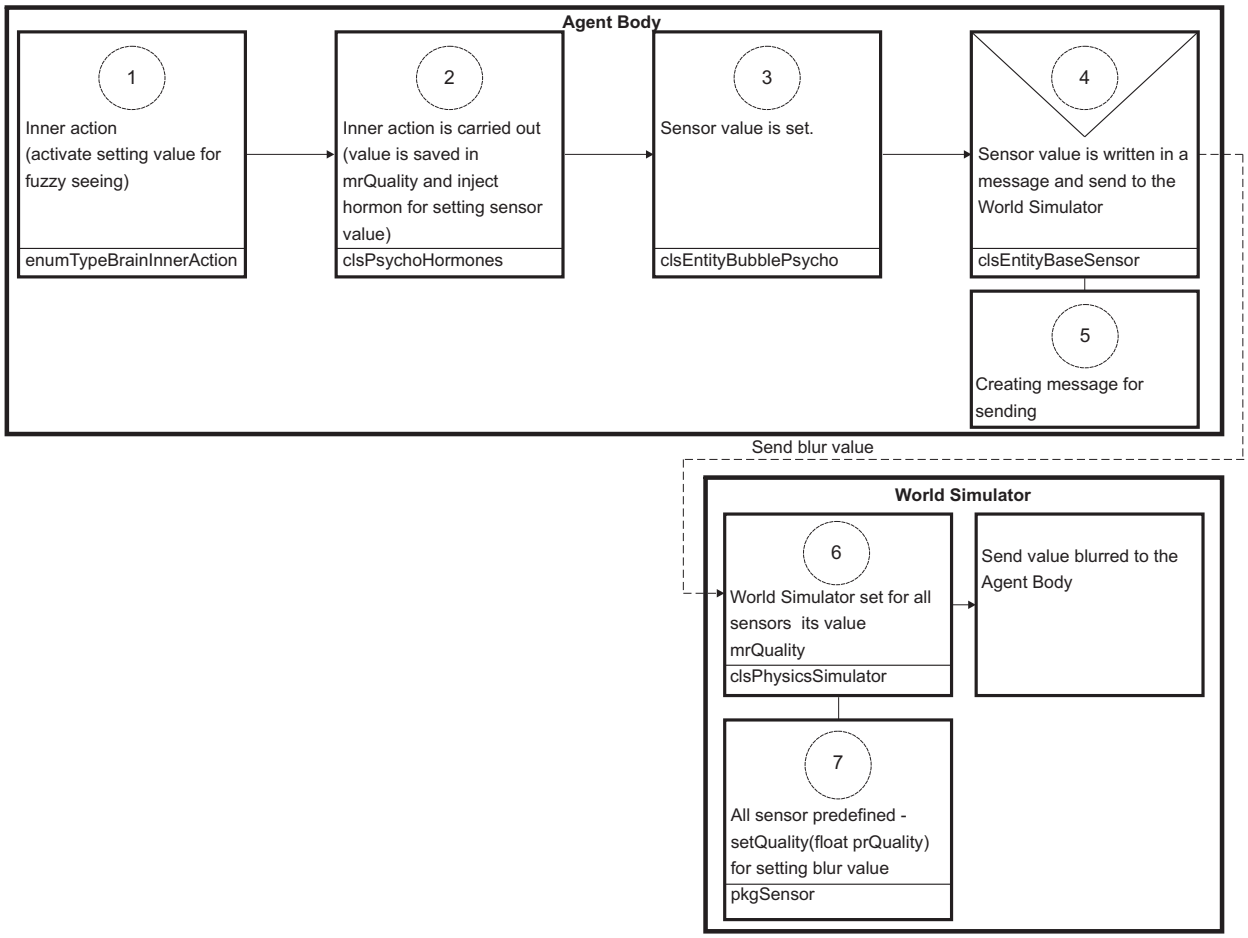


Figure 4.21: Information Flow of Setting Quality Value

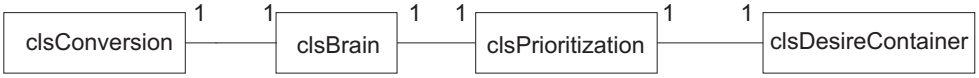


Figure 4.22: A Simplified Class Diagram for Conversion and Prioritization

In Figure 4.22 a simplified class diagram for the implementation of Conversion and the prioritization can be seen. As shown in Figure 4.22 the class *clsConversion* has an one to one relation to the *clsBrain*. It is situated in the function *doThinking* (cycle of it is explained in detail in Section 4.1) and in the variable moConversion of the class *clsBrain* all possible scenarios, which can activate the defense mechanism Conversion, are saved. In the implementation only one scenario can activate the defense mechanism Conversion, but in the file various conditions are possible for full recognition. By calling the function recognition it is checked if the scenario is full recognized, which has as result the setting of the variable *mnrecognized* to true and an inner action is loaded in the container. As long as the calculation of the fuzzy value is under one, the variable *mnrecognized* is true and activates every cycle the calculation of the value, which is done in the function

blur_calculation. If the value is equal or higher then one, the variable *mnrecognized* is set false and the function *recognized* is called until the scenario is full recognized again.

4.2.5 Prioritization of Defense Mechanisms

It is possible that two different defense mechanisms are initiated at the same time, wherefore a prioritization is necessary. By the realization of the four defense mechanisms, it is analyzed that three prioritizations are essentially, because Sublimation, Displacement and Intellectualization has an active action as result, but Conversion only initiates an inner action. An inner action is always executed and has no influence on outer actions, which are first sorted by priority before the action with the highest priority is set. So only when two or three of them are activated at the same time, a 'problem' occurs. In the following the reason for prioritization and solution for the problem is explained in detail.

Prioritization of Displacement and Sublimation

The prioritization is important, when Sublimation and Displacement are evoked at the same time. This can happen, because both reduce the level of the emotion rage. The difference between them is that before the Displacement is used a specific order of situations had to happen but Sublimation only has as condition the level of the emotion rage, which is the second condition for using Displacement. To minimize the possibility, that this situation happens, the threshold of the level of the emotion rage is for Displacement smaller than for Sublimation. But it can happen that the level of the emotion rage rises faster than Displacement can reduce it or it takes too long time before Displacement is used. Therefor it is essential to deactivate one e.g. by using the defense mechanism Repression. So one of them is repressed and the other one is utilized e.g. Displacement first, but if it takes too long, the other one becomes activated.

Prioritization of Displacement and Intellectualization

As well prioritization is important, when Displacement and Intellectualization are evoked at the same time. This results in a conflict because if a bubble is very angry and wants to attack another bubble, which is from the opposite team. But if there are too many enemies, which can help its teammate defending, it can be too dangerous and may end in death. So there also should be found a solution, e.g. repression of one.

Prioritization of Intellectualization and Sublimation

The prioritization is also important, when Intellectualization and Sublimation are evoked at the same time. The result is a conflict, because two actions, which cannot be done at the same time, are activated. So it has to be decided, which one is more important. The reason, why the chosen one is better at this moment, is not so easy to find, because the actions are not contrary like in the prioritization of Displacement and Intellectualization. It is possible to dance away, which would mean that both are used, because the bubble would dance and would go away from the enemies. Though this is not feasible to implement this in the BFG, wherefore also the defense mechanism Repression is used.

Solution for Prioritization

The prioritization is done with the functions *recognitionDefense* and *selectDefense* in the java class clsPrioritization. The function *recognitionDefense* checks, which defense mechanism is initiated. If only one is activated, nothing has to be done. If two or three are initiated, different solutions are implemented, which are included in the function *selectDefense*. The solution for the Displacement-Sublimation problem is that Displacement has the higher prioritization against Sublimation. But if Displacement takes too long to reduce the level of the emotion rage, it becomes deactivated and Sublimation is initiated. This can be seen as a Repression of Sublimation, because the bubble does not know that it ever wants to use Sublimation at the same time as Displacement. The solution for the Displacement-Intellectualization is depending on the levels of the emotions rage and fear. If the bubble has more fear than rage, it will prefer to flee, which means it uses Intellectualization. But if the bubble has more rage than fear, it will prefer to reduce the level of rage, which means it uses Displacement. This can also be seen as a repression of one of these mechanisms. If the level of rage and fear are equal, different solutions are possible, e.g. the random number, which is Gaussian variable, is splitted in two intervals and if the variable is smaller than 0.5 Intellectualization is used and if equal or greater than 0.5 it is used Displacement. But in the implementation Displacement is used if both levels are equal. The same solution is chosen for the Intellectualization-Sublimation problem.

As seen in Figure 4.22, the class *clsPrioritization* has a one to one relation to the class *clsDesireContainer*. In the class *clsDesireContainer* the full recognized scenarios activate the desires and defense mechanisms. Before this is done, it is checked in the class *clsPrioritization*, if prioritization is needed. Therefor the function *recognizedDefense* is responsible, i.e. it is controlled if only one defense mechanism is initiated. After that, the function *selectDefense* is called, which has as return value all Ids of the defense mechanisms, which should not be activated. These Ids are saved in an integer array called *index* and are returned back to the class *clsDesireContainer*. If no prioritization is needed, all values of the array *index* are null. In the class *clsDesireContainer* the values of the variable *index* is used to avoid an initiation of the defense mechanism, i.e. the defense mechanism is not loaded in the *oChangedScenarios* container. As seen in Figure **??** the class *clsPrioritization* has also a one to one relation to the class *clsBrain*. The reason therefor is that the variable *moPrioritization* is a parameter for triggering the recognition of Abstract Images.

4.2.6 An XML File Editor for Defense Mechanisms

To simplify creating and editing XML files of defense mechanisms a tool, shown in Figure 4.23, has been developed. As shown in Figure 4.23 the start window is splitted into two parts. On the left side the different files are listed as a tree. For example the "DEFENSE_ TEST3" is the name of a team as well as the name of the folder in which all files of this team are saved. This structure was already defined in the project ARS. As seen in Figure 4.23, the editor reads the files from the Abstract Image, Scenario and DefenseMechanism folder. When a file is selected the Abstact Images, Scenarios or Defense Mechanisms of one file are displayed in a table by the three characteristical variables ID, name and Description. The table is placed on the right side of the start window.

As seen in Figure 4.24 two different menu entries exist. By the first one the path for the XML file "GlobalConfig" is set. In this file the main path for the XML file "FactoryConfig"is written. In the XML file "FactoryConfig" the configuration of the teams, which are listed in the main

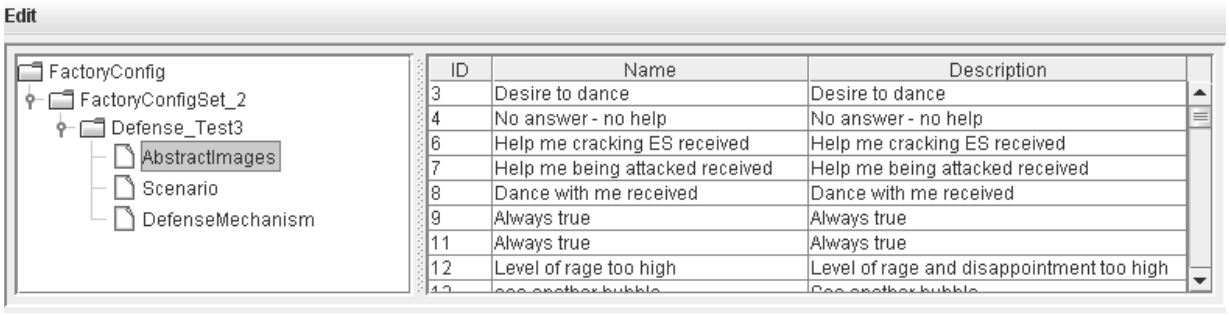


Figure 4.23: Editor

window on the left side, are written. The path in the XML file "GlobalConfig" is also the root directory for the files of the Abstract Images, Scenarios and Defense Mechanisms. It is used in the program to find the files for reading the content. To be able to read the files, the team name which is in the "FactoryConfig" file has to be append.

The second menu named newFile is for being able to create new XML files. Because every XML file is different in its design three submenus are created. At the beginning a file is loaded which contains the structure of one Abstract Image, Scenario or Defense Mechanism. For example in the new window it is possible to add nodes to create an Image, which is used to check the level of an emotion. To save this Abstract Image, Scenario or Defense Mechanism, it is possible to save this in an existing file as well as in a new file. This can be selected by the menu 'edit-save-save in existing file' or 'edit-save-save in new file'.

Figure 4.24: Editor menue

To see all informations of an Abstract Image, Scenario or Defense Mechanism, it is possible to select one by clicking on the desired one in the table. This causes the opening of a new window. At that point following two possibilities exist:

1. If an Abstract Image is selected, the opened window displays it in form of a tree. The tree can be expanded, to see all subnodes. The user can edit the values of the attributes of the Abstract Image by clicking on the desired node. Now the attributes of the node are listed in a table in the same window. If the changes should be saved, the menu 'edit-save' has to be chosen. It is possible to select the place where the file is saved as well as the name of the file. The window of the tree and the table of the attribute is seen in Figure 4.25.
2. When a Scenario or a Defense Mechanism is selected, the opened window shows a graph which demonstrates the state chart of the scenario or mechanism. In this window it is possible to select a state which causes the opening of a new window. In the new window the state is shown as tree and gives the user the possibility to edit the attributes of the state. The graph window is seen in Figure 4.26. The tree window of the state is the same

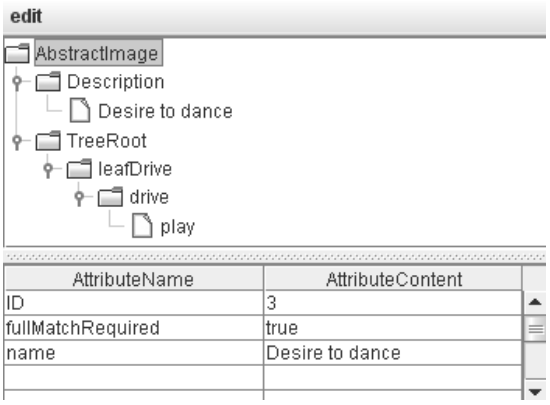


Figure 4.25: Editor Tree Abstract Image

as the window of the Abstract Image which is seen in Figure 4.25. In the graph the ID of the state is written in the circle which represents the state. On the arrow the ID of the Abstract Image which causes a state change is written. In this window it is only possible to edit the attributes of the states as well as the Abstract Image ID but not the attributes of the Abstract Images. This can be done by selecting the Abstract Image in the table of the main window.

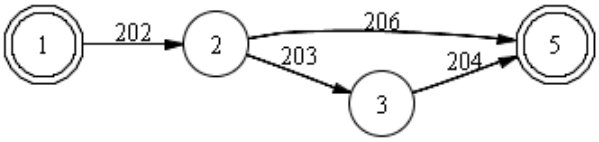


Figure 4.26: State Chart of a Scenario

To be able to show the XML files as tree the Document Object Model (DOM) of Java [1] is used. DOM allows direct access to the content of the XML file by loading the content into a tree structure of Java classes. It has various node types, whereas the following are the most important ones:

- A document node represents the whole tree structure.
- A document fragment node is a part of the tree structure.
- An element node represents exactly an element of HTML or XML.
- An attribute node represents exactly one attribute of HTML or XML. An code example in XML for an attribute is the following line, whereas the ID and the DirName are the attributes of the node named Team: `<Team ID="0" DirName="DEFENSE_ TEST3">`
- A text node represents the textual content of an element or attribute. An code example in XML for a text node is the following line, whereas Psycho becomes a text node which will be a sub node: `<Name>Psycho<Name>`

To be able to draw a graph like the state chart shown in Figure 4.26 two components have been used: the grappa library programmed in Java and the dot program. The dot program can be

[1] http://java.sun.com/j2se/1.4.2/docs/api/org/w3c/dom/package-summary.html

described following: “dot makes a hierarchical or layered drawings of directed graphs. The layout algorithm aims edges in the same direction (top to bottom, or left to right) and then attempts to avoid edge crossings and reduce edge length.”[2] The grappa library is responsible to visualize the graph which is created by the dot program. To be able to use the dot program a file, which includes all states and transitions of the graph, has to be generated. Afterwards the dot program transforms this file into a new file which can be read by the grappa library. Now this library can visualize the graph. The implementation of this in a Java program is done by the following steps:

1. First a text file is created, containing all states and transitions as well as informations like the label. This has to have a certain format which is seen in Figure 4.27. It is seen that every state has to be defined and it is possible to give them a *label*. This string is used for visualizing. The variable *tip* can be used to show more details in the tool tip when the cursor is over the state. The transition is defined by the string: starting state -> end state. The order of the statements is not predefined, i.e. the transitions of one state can be written before the other end states are defined. This is also seen in Figure 4.27.

```
digraph "Krise" {
rankdir = LR
1 [label="1", tip="InitState"];
1 -> 2[ label="202"];
2 [label="2", tip="decision if i want to dance"];
2[shape=circle];
2 -> 5[ label="206"];
2 -> 3[ label="203"];
3 [label="3", tip="i realy dont want to dance, but if i have a low social level - i will do it"];
3[shape=circle];
3 -> 5[ label="204"];
5 [label="5", tip="TerminalState"];
1[shape=doublecircle];5[shape=doublecircle];}
```

Figure 4.27: File before the dot program is used

2. After that a commando has to be executed which transforms this text file into another file. This new file can be used by the grappa library to draw the graph. The commando looks like following:

   ```
   dot.exe [-Tdot] InputFileName [-o] [OutputFileName]
   ```

   The commando is structured into five parts. The first part is dot.exe file. The second part is the format of the outputfile, whereas the -Tdot commando stands for default output format. The third part is the name of the input file, which is generated by the user in the first step. The commando “-o” - the fourth part - is responsible for using the written name as outputfile, which is the last part of the commando. This file is generated by transformation the input file and is used to draw the graph, i.e. this file is the input for the grappa library. It is important that both file names - the Input and the Output file - are specified by the whole paths because it is necessary to know exactly where they are saved. The design of this file is shown in Figure 4.28. In this file it is determined at which position the states are placed, which means that the dot program generates the graph. Finally the grappa library visualize the graph.

[2] http://www.graphviz.org/

```
digraph Krise {
        graph [rankdir=LR];
        node [label="\N"];
        graph [bb="0,0,328,73"];
        1 [label=1, tip=InitState, shape=doublecircle, pos="22,51", width="0.61", height="0.61"];
        2 [label=2, tip="decision if i want to dance", shape=circle, pos="118,51", width="0.50", height="0.50"];
        5 [label=5, tip=TerminalState, shape=doublecircle, pos="306,51", width="0.61", height="0.61"];
        3 [label=3, tip="i realy dont want to dance", shape=circle, pos="210,18", width="0.50", height="0.50"];
        1 -> 2 [label=202, pos="e,100,51 44,51 58,51 75,51 90,51", lp="72,58"];
        2 -> 5 [label=206, pos="e,284,53 136,54 142,55 148,56 154,57 195,60 243,57 274,54", lp="210,64"];
        2 -> 3 [label=203, pos="e,193,24 135,45 149,40 168,33 183,27", lp="164,45"];
        3 -> 5 [label=204, pos="e,285,44 227,24 240,28 259,35 275,41", lp="256,42"];
}
```

Figure 4.28: File after the dot program

4.3 Test Setups and Simulation

To test the functionality and the prioritization of the four defense mechanisms two different testbenches are developed and explained in the following. Finally a test in the simulator BFG is done, where the bubbles are equipped with all four defense mechanisms. The difference between the testbenches and the simulator BFG is that in the former, which are explained in Section 4.3.1 and 4.3.2, only one bubble is simulated and the recognition of the required Abstract Images are generated to simulate the defense mechanisms without being depended on the appearance of the needed situations. In the second one, which is described in Section 4.3.3, two different teams interact with the environment and each other and the decision unit is equipped with all four defense mechanisms to use one of them in an adequate situation.

4.3.1 Functionality Testbench

As seen in Figure 4.29 the Functionality Testbench consists of six parts. In the top the state chart of the scenario, which initiates the defense mechanism, and the state chart of the defense mechanism itself is shown. On the left side the complex emotions, in the middle right the blur value for the sensors and at the bottom right the basic emotions are drawn in response of time. In the middle the actions done by the bubble are also seen. The testbench includes those six parts, to visualize the active state, the trend of the levels of emotions, which is influenced by inner actions as well as by changes of levels of the complex emotions, and the outer actions, which are done by the bubble. The diagram of the blur value is only interesting for the defense mechanism Conversion.

Sublimation

The defense mechanism Sublimation is used to reduce the level of rage by dancing. If the level of rage is equal or over 5 (this is equal to the level MediumHigh and is also the defined threshold), the bubble starts to dance and if the level falls under a predefined threshold, it stops this action.

Intellectualization

The defense mechanism Intellectualization is used when too many enemies are attacking a teammate and the bubble flees instead to help. This action is caused because the bubble has too much fear. The reason for the bubble's behavior is that itself would die if it would help because there

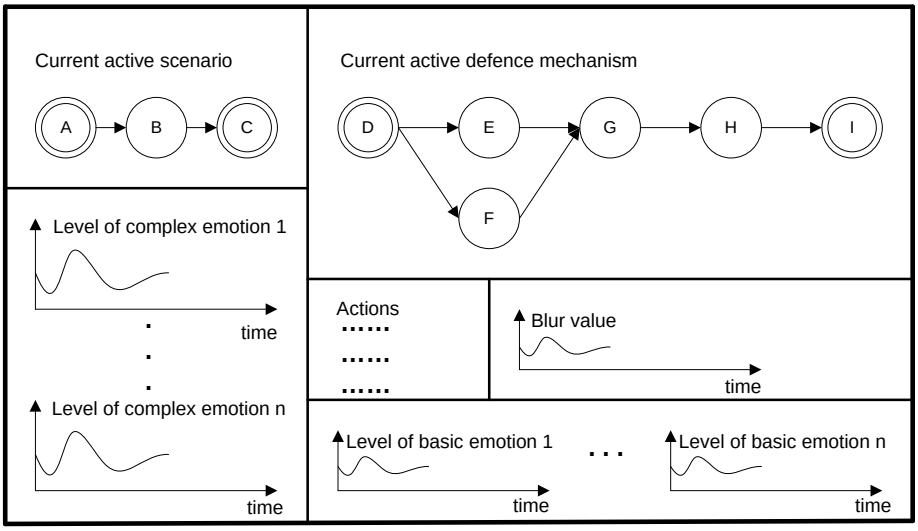


Figure 4.29: Functionality Testbench

are too many enemies or the other is too strong to have a chance to win. So the bubble would flee if the level of the emotion fear rises a defined threshold.

Displacement

There are two different reasons for using the defense mechanism Displacement. The first one is that nobody wants to help a bubble which needs someone for 'cracking' the energy source. The second one is that a bubble wants to dance with someone but nobody wants to dance with it. A bubble has two different possibility to discharge the inner tension. The first one is not helping someone who needs help e.g. for defense, the second one is to attack a bubble of the opposite team. The defense mechanism Displacement is used to reduce the level of rage, which rises when one of the reasons explained above had happened. It does not matter how the discharging is done because both possibilities do it with the same intensity.

Conversion

The defense mechanism Conversion has only an impact on the sensor, which is concerned with the seeing of objects in the world. This sensor gives the real position of an object if the blur value is one and an object is not seen if the blur value is zero. So the blur value can only be in the range between zero and one. To calculate the blur value the level of the emotions hope, joy, disappointment and rage are included as well as the dependence of time. The dependence of the time is done because of giving the possibility to deal with Conversion as it can be done in the real world.

As seen in the paragraphs above, the defense mechanisms Displacement and Conversion are influenced by the level of emotions, wherefore they are tested within various start conditions, which means the level of all emotions are varying at the beginning in different test cases. Because of also starting with very low levels of emotion, it can be necessary that the recognition of the scenarios initiating the defense mechanisms has to happen more than once. Also interesting is, how the abortion conditions influences the defense mechanism Displacement.

4.3.2 Prioritization Testbench

For testing the prioritization of the four defense mechanisms, different situations are used. As explained in Section 4.2.5, three different prioritizations are implemented. The testbench for these three test cases is shown in Figure 4.30, which shows four state charts - two scenario and two defense mechanism state charts. If both scenarios are full recognized at the same time, two defense mechanisms are initiated. Because not be able to executed both at the same time, one should be canceled and the other one should be completed. To see the activated state of the state charts, they are highlighted in another color as the other state. If the executed one could not solve the problem (inner conflict), the canceled one becomes activated after deactivating the first tried one. To see the influence of this abortion on the levels of the emotions, they are also displayed in the testbench.

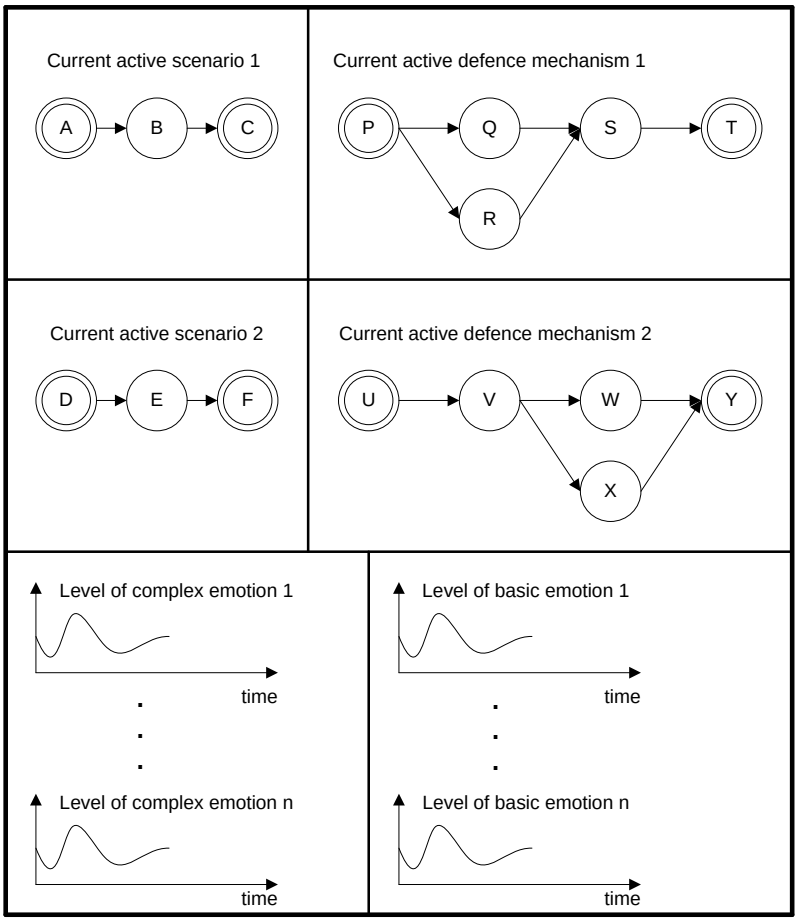


Figure 4.30: Prioritization Testbench

4.3.3 Simulator

In the simulator BFG explained in detail in Section 2.2, two different teams are interacting with the environment and with its teammates to fulfill their desires and to survive as long as possible. One team has the defense mechanisms as knowledge base and the other has not. The team, which has no defense mechanism, is utilized to make it possible that the defense mechanism

Displacement can be used also when another bubble of the opposite team is seen. The other reason for the second team is that one team is only used to provoke the inner conflict, e.g. for Intellectualization. The simulation should show, if the bubbles use the defense mechanisms in special situations. For testing two different simulations are done:

1. In the first simulation the bubbles of both teams have the XML files for being able to satisfy the desire to dance as well as the second team has the XML files of the defense mechanisms. In this simulation, the defense mechanism Intellectualization may occur not so often, because the bubble has more prioritization for fulfilling the desire to dance. Thus they may be more frustrated if this happens not so often and therefor use the defense mechanisms Sublimation, Displacement and Conversion. In this simulation the bubble only has the XML files of the defense mechanism Displacement for solving the problem that another teammate would not like to dance with it. This is done, because the bubble has only the XML files for satisfying the desire to dance and not for satisfying the desire to eat. Thus the defense mechanism Displacement for solving the inner conflict that occurs because of nobody would like to help 'cracking' the energy source would always be activated.
2. In the second simulation the bubble has the XML files for being able to satisfy the desire to eat, but it is important that the energy sources can only be 'cracked' in teams of at least two bubbles, as well as the second team has the XML files of the defense mechanisms. In this simulation, the defense mechanism Intellectualization may occur not so often, because the bubble has more prioritization for fulfilling the desire to eat. Thus they may be more frustrated if this happens not so often and therefor use the defense mechanisms Sublimation, Displacement and Conversion. As well in this simulation the bubble has only the XML files of the defense mechanism Displacement for solving the problem that another teammate would not like to help 'cracking' the energy source. This is done because the bubble has only the XML files for satisfying the desire to eat and not for satisfying the desire to dance. Thus the defense mechanism Displacement for solving the inner conflict that occurs because of nobody would like to dance would always be activated.

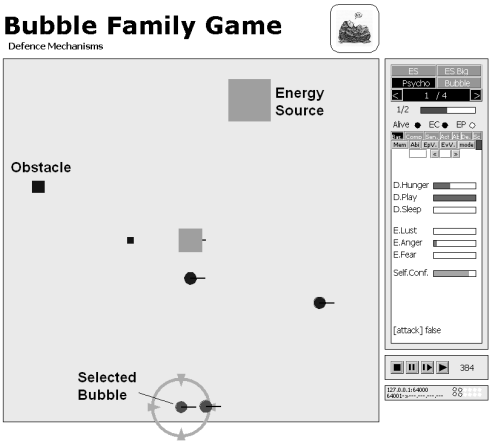


Figure 4.31: The Simulator BFG

In Figure 4.31 the simulator BFG is shown. The bubbles of one team are in the same color and are drawn as circles. The quadrate in light grey are the energy sources and the quadrate in dark grey are obstacle. On the left side the environment of the bubbles is shown and on the right

side different values of the chosen bubble are displayed. There the level of the basic and complex emotions, the blur value and the initiated scenarios and desire are demonstrated. So it can be seen, if a bubble uses the defense mechanisms or not.

In this chapter the technical realization of the four defense mechanisms Sublimation, Intellectualization, Displacement and Conversion is explained in detail. Therefor some basics like the scenario recognition system and the function *doThinking* are described first. Afterwards a general cycle of the defense mechanisms is shown to demonstrate how they work. Then an example and the implementation of them in the BFG is explained. Furthermore the reason for the prioritization, which is needed when two defense mechanisms are initiated at the same time, and the solution of the problem is presented. Finally the testbenches for functionality and prioritization and the simulation is described. In the next chapter the results of these tests are shown.

5 Simulation Results

In this chapter the simulation results are presented. First the results of the functionality tests under different start conditions are shown. Afterwards the outcome of the prioritization tests is described. Finally the results of the simulation in the Bubble Family Game (BFG) are explained. In the test environment of the functionality and the prioritization test the perception of an image is controlled by clicking on a button. If it is clicked on the button the bubble perceives a new image. The time, how long a bubble uses the defense mechanism, is compared by the number of clicks, which are dependent on the start conditions. The usage of a defense mechanism starts when the associated scenario is completely recognized, i.e. since the bubble receives the first image.

5.1 Functionality Tests

As presented in Section 4.3.1 there are four functionality tests. Because of being dependent on the levels of emotion, different start conditions are chosen to test how long the bubble would use the defense mechanism. Because of being dependent on the defense mechanism which start conditions are interesting, they are described in detail for every test case.

Sublimation

For testing the functionality of Sublimation the following different situations are used, whereas at all tests the level of the emotion rage is set to MediumHigh(equal to 5)/High(equal to 6)/Very-High(equal to 7) and the level of the emotion disappointment is also set to high (equal to 4) at the beginning:

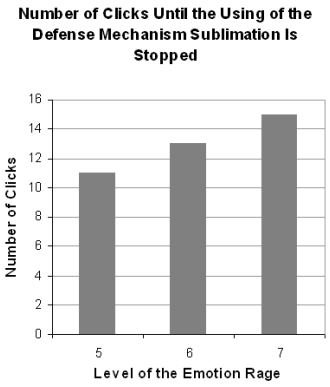
1. First only the required images for the scenario and the defense mechanism are recognized. So no influences of other recognized images as well as no influences of changes in the level of other emotions are checked, which could influence the time of using the defense mechanism Sublimation.
2. Second the changes in the level of other emotions are done, but only to improve the usage of the defense mechanism Sublimation. So the level of the emotion hope rises continuous, which has as result that the emotion rage falls.

3. Finally changes in the level of other emotions are done, but only to degrade the usage of the defense mechanism Sublimation. So the level of the emotion hope falls continuous, which has as result that the emotion rage rises.

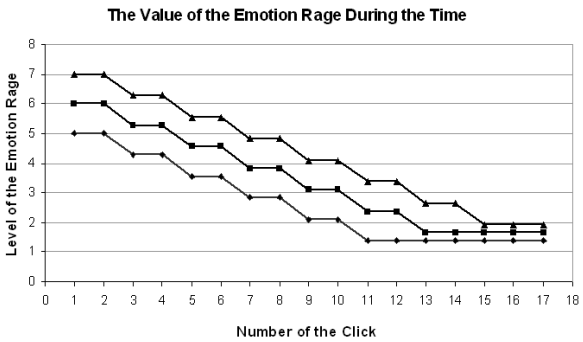
The choice of the start level of the emotion rage is done because Sublimation becomes activated at a level greater or equal MediumHigh. The reducing of the level of the emotion rage is done with an inner action with the constant value of 0.3, whereas the value can be between 0 and 1.

The high of the influence on the level of basic emotions by changes of other complex emotions is written in the file, where defined is, which complex emotions are available. In this file a constant factor is written, which can be between 0 and 1. This value is used for calculating another value, which is exponential distributed. The calculated value is the value change of the emotion.

In the following the results of the first test condition are presented. As seen in Figure 5.1(a) the time of using the defense mechanism Sublimation rises when the start level of the emotion rage becomes higher. Because of reducing the level of the emotion rage with the constant value of 0.3, the time of using the defense mechanism rises also constantly. In Figure 5.1(b) the course of the level emotion rage is shown, whereas the start conditions has been changed. The level of the emotion falls only after two clicks are done. This is the result of the way of implementation. At the first click the bubble receives the first image, which is used to check if the level is over the threshold. Because of exceeding the threshold the bubble starts dancing which results in an inner action that reduces the level. This can be seen in Figure 5.1(b) by the falling line between the second and third click. At the third click it is checked if the level is over or under the threshold and at that time no inner action is set. The result is that the level of the emotion is the same value as before. If the level is still over the threshold the bubble dances further, otherwise the using of the defense mechanism is stopped.



(a) Time of using Sublimation

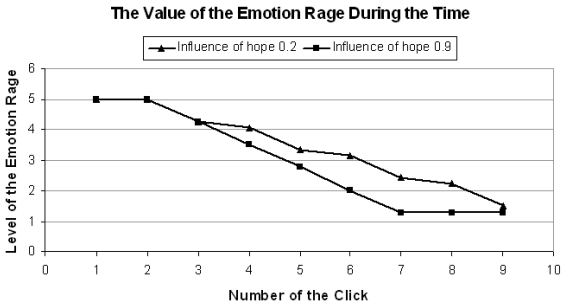


(b) The Course of the Level of the Emotion Rage

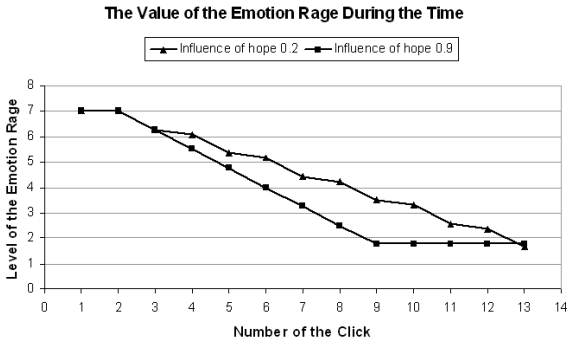
Figure 5.1: Results of the First Test Condition

In the following the results of the second test condition are presented. For this test the level of the emotion rage is set to MediumHigh (equal to 5)/VeryHigh (equal to 7). The value for the influence of the emotion hope is set to 0.2 and to 0.9, whereas the first means a small and the second a high influence on the level of the emotion rage. The change of the level of the emotion rage is only done if the level of the emotion hope changes. As seen in Figure 5.2(a) the level of emotion rage changes by every click. This is caused by the changing of the level of the emotion hope at the same time as the level of the emotion rage is checked (no inner action is set at that time). In this test case the level of hope rises. This is shown in Figures 5.2(a) and 5.2(b) because

the level of the emotion rage falls. In comparison to the course seen in Figure 5.1(b) the level of emotion is faster under the threshold.



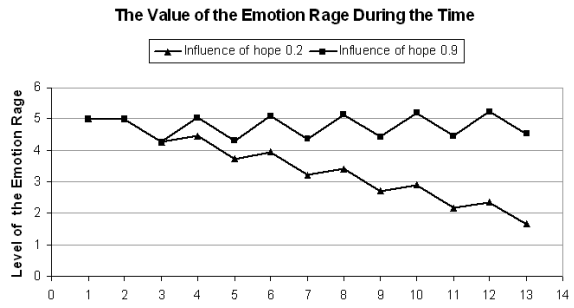
(a) The Course of the Level of the Emotion Rage Influence by Hope



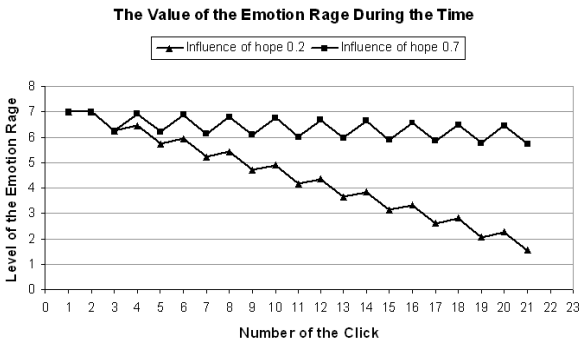
(b) The Course of the Level of the Emotion Rage Influence by Hope

Figure 5.2: Results of the Second Test Condition

In the following the results of the third test condition are presented. For this test the level of the emotion rage is set to MediumHigh (equal to 5)/VeryHigh (equal to 7). The value for the influence of the emotion hope is set first to 0.2 and afterwards either to 0.9 or 0.7. As seen in Figure 5.3(a) the level of emotion rage changes by every click. This is caused by the changing of the level of the emotion hope at the same time as the level of the emotion rage is checked (no inner action is set at that time). In this test case the level of hope falls. This is seen in the Figures 5.3(a) and 5.3(b) because the level of the emotion rage rises when the level of the emotion hope falls. That is the reason why the course always changes the direction - if the bubble dance the level of the emotion rage falls but if hope falls the level of the emotion rage rises. In Figure 5.2(a) it is seen that the value of the influence is set to 0.9, which influences the level of the emotion rage so high that it is not possible to reduce the it. This is the result of the high influence of the emotion hope corresponding to the emotion rage. Because of influencing the level of the emotion too much the second test case seen in Figure 5.3(b) is done by an influence of 0.7.



(a) The Course of the Level of the Emotion Rage Influence by Hope



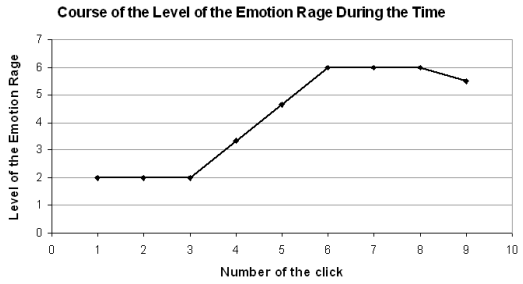
(b) The Course of the Level of the Emotion Rage Influence by Hope

Figure 5.3: Results of the Third Test Condition

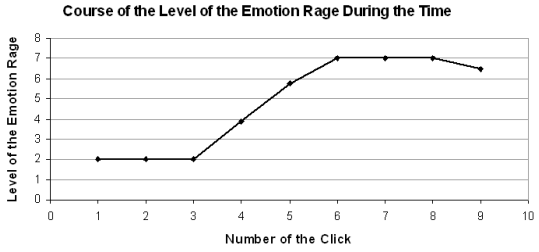
Displacement

For testing the functionality of Displacement it is only interesting if the mechanism is used if the level of the emotion rage exceeds the predefined threshold. If the mechanism is utilized the level of the emotion rage is reduce by an inner action, whereas the value for reducing is predefined in the XML file. For testing, the level of the emotion rage is set to Low (equal to 2) at the beginning. The threshold for using the defense mechanism Displacement is set to Medium (equal to 4). As explained in Section 4.3.1, two different possibilities of using the defense mechanism exist. The first one is that it is not possible to fulfill the desire to dance, the second one is that it is not possible to fulfill desire to eat. That is the reason why two test are done.

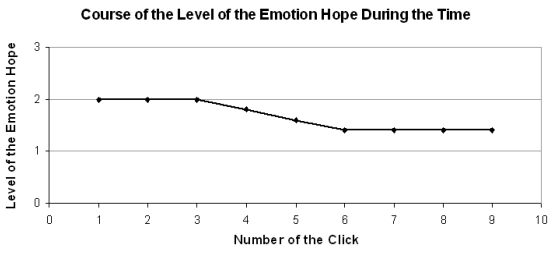
After starting the test the bubble has the desire to dance. If no one answer its question for dancing the level of the emotions rage rises and the level of the emotion hope falls. Due to the test conditions the exceeding of the threshold happens not at the first time where a bubble gets now answer. That is the reason why the desire has to be arise more than one time. If the threshold is exceeded the bubble uses the defense mechanism Displacement by attacking another bubble or not helping if another bubble needs help. By setting this action the level of the emotion falls. But is falls only once time because the inner action is set once time. The course of the level of the emotion rage is seen in Figure 5.4(a) and of the emotion hope is seen in Figure 5.4(c), whereas at that test case only the influence of the emotion hope is important because of this value changes during the time. The course of the level of the emotion rage is seen in Figure 5.4(b) and of the emotion hope, disappointment and pride are seen in Figure 5.4(d), whereas at that test case the influence of the emotion hope, disappointment and pride are important because of all of them change during the time. That test case "Not Fulfilling the Desire to Eat" yields the same results as the test case "Not Fulfilling the Desire to Dance".



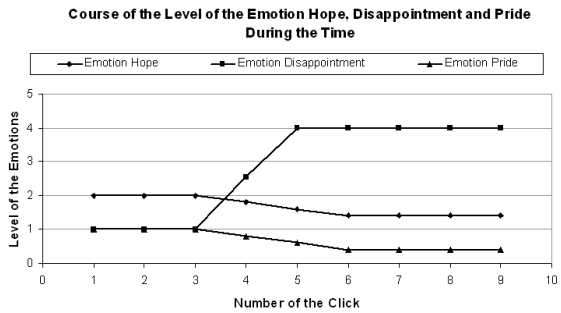
(a) The Course of the Level of the Emotion Rage



(b) The Course of the Level of the Emotion Hope



(c) The Course of the Level of the Emotion Rage



(d) The Course of the Level of the Emotion Hope, Disappointment and Pride

Figure 5.4: Results of the First Test

Intellectualization

For the testing of the functionality of Intellectualization it is only interesting if the mechanism is used if the level of the emotion fear exceeds the predefined threshold. To test this the level of emotion is under the threshold and rises continuously. The test result is that Intellectualization is only used when the level of the emotion fear exceeds the threshold.

Conversion

For testing the functionality of Conversion following different situations are used:

1. the level of the 'positive' emotions like hope are very low and the level of 'negative' emotions like rage are very high. In this case the levels of the desire to eat and play do not have to be very high to activate the defense mechanism. The result is the time the bubble needs to see all exactly. For this test the levels of the emotions are constant, i.e. they do not change over the time.
2. The level of the 'negative' emotions like disappointment are very low and the level of 'positive' emotions like joy are very high. In this case the levels of the desire to eat and play have to be very high to activate the defense mechanism. The result is the time the bubble needs to see all exactly. For this test the levels of the emotions are constant, i.e. they do not change over the time.
3. Both test described in 1. and 2. are done again. The difference to the other tests is that now levels of the emotions are not constant, i.e. they change over the time. The result is the time the bubble needs to see all exactly. As well it can be shown how the changes of the levels of emotions influences the time of seeing all at the right position.

For all graphs on the x axis the time difference between the start of using the defense mechanism Conversion is applied, i.e. it can be seen in the graph how long the bubble uses the mechanism. To calculate the time difference, the system time is used. This time is in milli seconds so to calculated the minutes the value is divided by 60.000. The unit of the time difference is minutes.

In the following the results of the first test condition are presented. As shown in Figure 5.5 the bubble uses the defense mechanism Conversion for 0.82 minutes. In that test case the level of the emotions rage was High (equal to 6), disappointment was also High (equal to 3), hope was normal (equal to 2) and joy was Low (equal to 1). That is the reason why the bubble needs so much time for seeing everything exactly. Because of the high levels of the 'negative' emotions and the low levels of the 'positive' emotions the level of the fuzziness starts at zero.

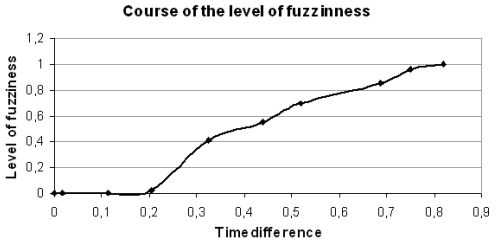


Figure 5.5: The Course of the Level of Fuzziness

In the following the results of the second test condition are presented. As seen in Figure 5.6 the bubble uses the defense mechanism Conversion approximately 0.43 minutes. In that test case the level of the emotion rage was Low (equal to 2), the level of the emotion disappointment was also Low (equal to 1) and the level of the emotions hope was VeryHigh (equal to 4)and joy was High (equal to 3). That is the reason why the bubble needs so less time for seeing everything exactly. Because of the low levels of the 'negative' emotions and the high levels of the 'positive' emotions the level of the fuzziness falls only to 0.7. In comparison to Figure 5.5 the bubble needs the half of time to see everything exactly and the level of the fuzziness is bigger, which means the bubbles sees less fuzzy, if the emotions of the bubble are positive.

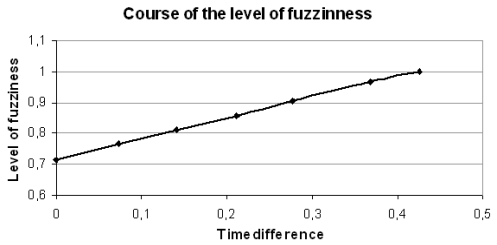
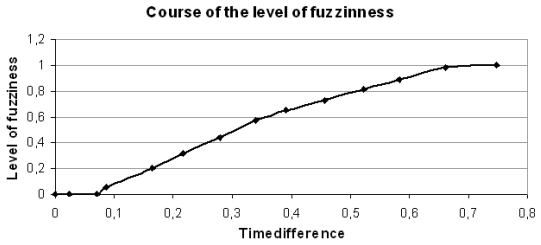
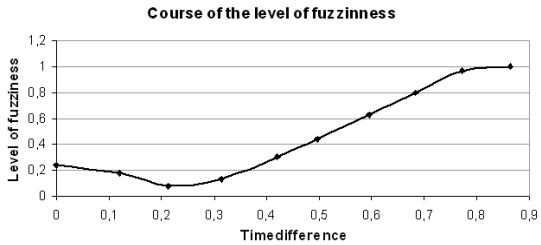


Figure 5.6: The Course of the Level of Fuzziness

In the following the results of the third test condition are presented. Figure 5.7(a) shows the course of the level of the fuzziness when the level of the emotion disappointment falls and the level of the emotions joy and hope rises. But the start level of the 'negative' emotions are high and the start level of the 'positive' emotions are low, which leads to the start level of the fuzziness zero. Figure 5.7(b) shows the course of the level of the fuzziness when the level of the emotion disappointment and hope rise but the level of the emotions joy falls. The start level of the 'negative' emotions are low and the start level of the 'positive' emotions are high, which leads to the start level of the fuzziness 0.24. Because of the high influence of the emotion disappointment on the emotion rage, the level of the emotion rage rises fast. This can be seen in Figure 5.7(b) by falling first. At a time of 0.21 the level of the emotions rage, disappointment and hope are maximum, which leads to the rising of the course of the level of the fuzziness. The reason therefor is the influence of the time because the longer the time is elapsed the smaller is the fuzziness of seeing.



(a) The Course of the Level of Fuzziness



(b) The Course of the Level of Fuzziness

Figure 5.7: Results of the Third Test

5.2 Prioritization Tests

As presented in Section 4.3.2 there are three situations where prioritization is needed and the prioritization depends either on the level of the emotions rage and fear or on the amount of usage. In the following the three test cases for the prioritization are explained in detail.

Displacement and Sublimation

In this case the reason of using one of the mechanisms is not the level of an emotion, because both of them are activated when the level of rage is too high. So it is decided that Displacement is activated at a lower level than Sublimation. This is done because Displacement needs more preconditions as Sublimation. This helps to avoid the situation that both are activated at the same time. But if the level rises in spite of using Displacement, this mechanism becomes deactivated and Sublimation becomes activated. Another situation where prioritization is needed is if the level of the emotion rage is over the threshold of Sublimation, because it is also possible that both mechanism become activated. So in this case it is tested if the defense mechanism Displacement becomes activated first if the threshold of Sublimation, which is higher then for Displacement, is exceeded. As well it is tested if it is deactivated when the second threshold is exceeded in the time of execution, i.e. Sublimation should be activated. The results of these tests are that Displacement is used first if the first threshold is exceeded and becomes deactivated if the abortion condition (exceeding the second threshold) is recognized. This activates the defense mechanism Sublimation because the level of the emotion rage is still over the threshold.

Displacement and Intellectualization

In this case the reason for using one of the mechanisms are the levels of the emotions fear and rage. It is tested what happens if the level of the emotion rage is higher than or equal to the level of the emotion fear as well as what happens if the level of emotions rage is lower than the level of the emotion fear. Also interesting is what happens if the abortion condition of the executed one is perceived. But only the defense mechanism Displacement has an abortion condition. So if Displacement becomes aborted Intellectualization is used if the level of the emotion fear is still high enough. The test results are that Displacement is used if the level of the emotion rage is higher than or equal to the level of the emotion fear. Intellectualization is used if the level of the emotion fear is higher than the level of the emotion rage. If Displacement is aborted because the level exceeds the second threshold, Intellectualization is used if the level of the emotion fear is still high enough. In this test case only the defense mechanism Displacement and Intellectualization are possible to be used because if the level of the abortion condition would exceed also the threshold of Sublimation another prioritization has to be used.

Intellectualization and Sublimation

In this case the reason for using one of the mechanisms are the levels of the emotions fear and rage. It is tested what happens if the level of the emotion rage is higher than or equal to the level of the emotion fear as well as what happens if the level of emotions rage is lower than the level of the emotion fear. The results of the test are that Sublimation is used if the level of the emotion rage is higher than or equal to the level of the emotion fear. Intellectualization is used if the level of the emotion fear is higher than the level of the emotion rage.

5.3 Simulation

As presented in Section 4.3.3 two tests are done. To be able to make these tests various XML files have to be prepared. In the following the different test conditions are explained in detail. It is described, which XML files are used, how many bubbles are interacting in the environment and the chronology of the scenario.

Simulation 'Desire to dance'

The first test is done with XML files, which enables the bubble to fulfill the desire to dance. The test is done by looking at the cycle of the defense mechanism Displacement for the desire to dance. Three states of the flow chart shown in Section 4.2.3 are chosen to see if the defense mechanism is used. The first state is state number 2. In this state the bubble had asked if another one would like to dance with it and waits for an answer. This is shown in Figure 5.8.

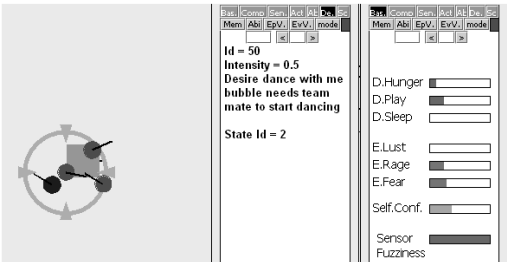


Figure 5.8: Results of the Simulation 'Desire to dance' - State 2

The second state is state number 4. In this state the bubble had get no answer and the level of the emotion rage rises. The change can be seen by comparing the level of the emotion in Figure 5.8 and 5.9.

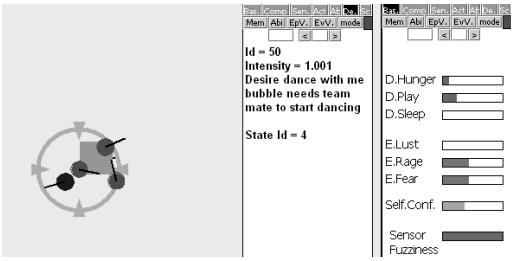


Figure 5.9: Results of the Simulation 'Desire to dance' - State 4

The third and last state is state number 7. If the defense mechanism is used until the end, the listed file is removed from the list seen in Figure 5.10 on the left side. After that has happened the level of the emotion rage falls. This change can be seen by comparing the level of the emotion in Figure 5.9 and 5.10. The Figures below shows the list of the desire and defense mechanisms, which are activated as well as the levels of the drives and basic emotions. In this test case three bubble are in the first team, which are able to use defense mechanisms and only one bubble is in the opposite team, to enable the use of Displacement by attacking.

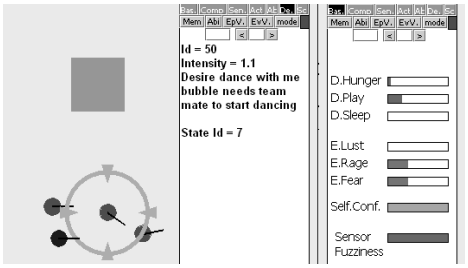


Figure 5.10: Results of the Simulation 'Desire to dance' - State 7

Simulation 'Desire to eat'

The second test is done with XML files, which enables the bubble to fulfill the desire to eat. The test is done by looking at the cycle of the defense mechanism Displacement for the desire to eat. Three states of the flow chart seen in Section 4.2.3 are chosen to see if the defense mechanism is used. The first state is state number 2. In this state the bubble had asked if another one would like to help it 'cracking' the energy source and waits for an answer. This is shown in Figure 5.11.

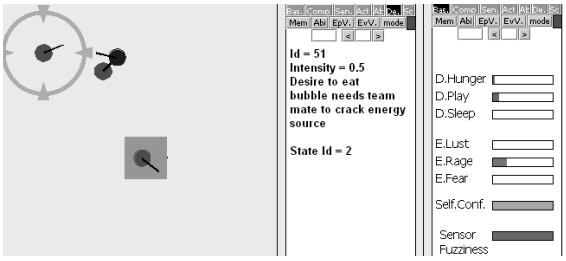


Figure 5.11: Results of the Simulation 'Desire to eat' - State 2

The second state is state number 6. In this state the bubble had get no answer and the level of the emotion rage rises. The change can be seen by comparing the level of the emotion in Figure 5.11 and 5.12.

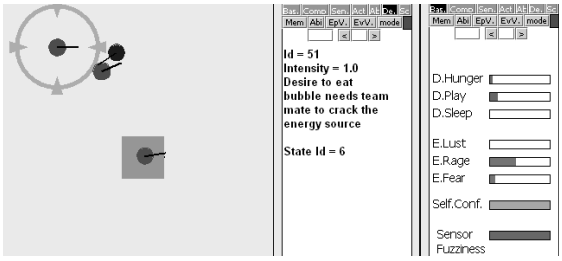


Figure 5.12: Results of the Simulation 'Desire to eat' - State 6

The third and last state is state number 9. If the defense mechanism is used until the end, the listed file is removed from the list seen in Figure **??** on the left side. After that has happened the level of the emotion rage falls. This change is seen by comparing the level of the emotion in Figure 5.12 and 5.13. The Figures below shows the list of the desire and defense mechanisms, which are activated as well as the levels of the drives and basic emotions. In this test case three bubble are in the first team, which are able to use defense mechanisms and only one bubble is in the opposite team, to enable the use of Displacement by attacking.

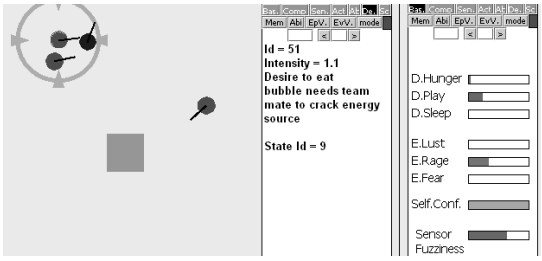


Figure 5.13: Results of the Simulation 'Desire to eat' - State 9

In this chapter the test results of the functionality tests, the prioritization test and the simulation in the BFG e presented. For the three tests different test situations and start conditions have chosen to show how the defense mechanisms works and how the bubbles have an advance by using them. In the next chapter a conclusion is presented and an outlook for further work is given.

6 Conclusion and Outlook

A challenge for AI (Artificial Intelligence) is that agents should be able to interact in an unknown environment without help. Therefor the agent has to receive information about the environment, has to make decisions and has to have the possibility to interact with its environment e.g. by hands. The focus of this work are defense mechanisms which are a part of 'Decision making'.

The first approach to improve the decision making process is done by the integration of emotions. As explained in Section 2.1, it has to be find definitions of them at the beginning, to be able to implement them in the decision unit of robots, because the concept of emotions is very broad. Furthermore various aspects including physiological, motivational and expressive ones can be covered. Various existing approaches differ significantly concerning to their aims and assumption as well as refer in several degrees to existing theories of emotions.

Another approach is used by A. Buller. He utilizes psychodynamics for designing his decision unit and implements it in robots like Miao-V. To improve activities as self-initiated exploration of an environment, he develops a controller named Volitron. In this controller he also implements defense mechanisms. Other researchers develops decision units by mixing theories like psychology and neurology, whereby they take only parts of the theory, which solves the upcoming problem best. But some of these theories are contradictory. The team of the project ARS (Artificial Recognition System) of the Institute of Computer Technology on the Vienna University of Technology, uses only one concept namely psychoanalysis, which is described in the following.

The team of the project ARS (Artificial Recognition System) decides to use psychoanalytical models especially the second topographical model of Sigmund Freud for designing a decision unit. In his model, Freud distinguished between the three instances Ego, Id and Superego, which are in constant contact with each other. The *Id* is responsible for the human drives and exists since birth. The *Superego* contains the forbiddances as well as the restrictions and demands of the parents and evolves during the childhood. The *Ego* is the connection to the outer world and tries to satisfy the drives of the *Id* by including the conditions of the outer world as well as the forbiddances and restrictions of the *Superego* in the decision making process. As well Freud had integrated concepts of emotions and feelings in his model, which influence the decision making process too.

A decision making process, which is influenced by drives or emotions, provokes other problems. If a drive would like to become satisfied, but the conditions of the outer world denies this, the drive can block other actions. But these actions might make it possible to satisfy the drive later because the conditions of the outer world have changed by them. This leads to frustration because

no satisfying action is executed. To avoid such situations psychoanalytical defense mechanisms are used.

The model of the project ARS is implemented in the simulator BFG (Bubble Family Game). In the BFG agents have to reach goals alone as well as in teams and their actions are influenced by drives and emotions. They have to survive as long as possible as well as to fulfill their drives. This simulator is used to implement the concept of defense mechanisms.

Psychoanalytical defense mechanisms are a function of the Ego and enables the Ego to distinguish between thoughts, which are allowed and which not like a censor. The forbidden drives can be repressed or distorted or the instinctual aim is displaced, suppressed or inhibited. To do this the Ego has various mechanisms allocated. In this work, defense mechanisms have been described in detail. As an example the defense mechanism Repression can be used to repress the forbidden drive, which means that the drive is unconscious. Another solution can be Intellectualization, which is used to set another action, which has a logical reason. For these psychoanalytical functions a technical model has been built and implemented in the BFG. Afterwards different tests like a functionality tests are done to show how the defense mechanisms works.

As explained in Section 3.1.2 defense mechanisms can be applied to solve different problems like deadlocks or endless lops with a new approach. Because of being used as support for perception and decision making, it is important that the decision unit of the agent is designed by approaches of psychoanalysis or at least psychodynamics when using the concepts of defense mechanism. This is necessary because defense mechanisms are utilized to help the agent if it is not possible to rebuilt the inner balance, which is disturbed by unfulfilled drives and wishes. This work shows that the application of defense mechanisms helps the agent to control negative emotions like rage as well as to facilitate the rebuilding of the inner balance. It is also shown that defense mechanisms can be used to repress defect sensors. Another result of this work is the prioritization of actions, plans, etc. which can be done by using defense mechanisms.

In further work the focus should be drawn to the research, if defense mechanisms have similarities, which enable them to assemble into groups. This is important, because it may be easier to realize defense mechanisms e.g. some of them need special information like to know who is involved in a specific situation where a defense mechanism is used. Another point, which has to be kept in mind is that the defense mechanisms must be prioritizied. This means that it has to be arranged which defense mechanism should be utilized in which situations and when another one should become activated instead of the first used.

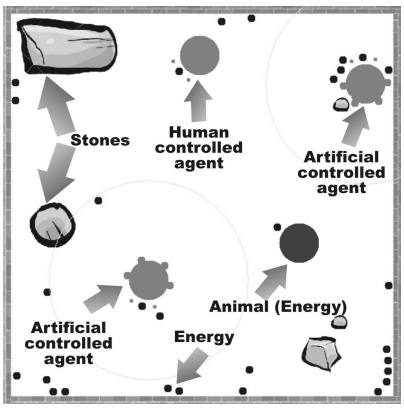


Figure 6.1: The New Simulator [ZLM09]

In further work it should also be concentrated on the research if it is meaningful to use all mechanisms in one agent. So it should be decided which defense mechanism is reasonable in the agent and which not. This work defines a basic framework for further research.

The team of the project ARS has begun to develop a second model, in which defense mechanisms will be integrated in further work, whereas the defense mechanisms would already have a key role in the perception. This is done because the perception of memory traces, which are described in detail in [ZLM09], become activated if they are perceived. But defense mechanisms should avoid this, if e.g. the drive should not become activated because of repressing. So they decide if a perception - drive as well as outer perception - becomes conscious or not.

As well this team has started to design a new simulator environment shown in Figure 6.1, which implements the new model by using Mason as development environment. In this new simulator, the model of the defense mechanisms will be integrated in further work.

Literature

[AWA04] Aronson, E. ; Wilson, T. D. ; Akert, R. M.: *Sozialpsychologie*. Pearson Studium, 2004

[BLPV07] Burgstaller, W. ; Lang, R. ; Pörscht, P. ; Velik, R.: Technical Model for Basic and Complex Emotions. In: *Proceedings of 2007 IEEE International Conference of Industrial Informatics* (2007), p. 1033–1038

[Bul02] Buller, A.: Volitron: On a Psychodynamic Robot and Its Four Realities. In: *Proceedings Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* 94 (2002), p. 17–20

[Bul05] Buller, A.: Building Brains for Robots: A Psychodynamic Approach. In: *Invited talk on the First International Conference on Pattern Recognition and Machine Intelligence, PReMIT'05* (2005), p. 17–20

[Can97] Canamero, D.: Modeling motivations and emotions as a basis for intelligent behaviour, 1997

[DFKU09] Dietrich, D. ; Fodor, G. ; Kastner, W. ; Ulieru, M.: Considering a Technical Realization of a Neuropsychoanalytical Model of the Mind – A Theoretical Framework. In: Dietrich, Dietmar (Hrsg.) ; Fodor, Georg (Hrsg.) ; Zucker, Gerhard (Hrsg.) ; Bruckner, Dietmar (Hrsg.): *Simulating the Mind – A Technical Neuropsychoanalytical Approach*, Springer, Wien, 2009. – ISBN 9781412949644, p. 99 – 115

[DFZB08] Dietrich, D. ; Fodor, G. ; Zucker, G. ; Bruckner, D.: *Simulating the mind*. Wien : Springer Verlag, 2008

[DTM+09] Deutsch, T. ; Tmej, A. ; Muchitsch, C. ; Zucker, G. ; Riediger, C. ; Lang, R.: Failsafe Aspects of a Decision Unit Inspired by Cognitive Sciences - The Id without Ego and Super-Ego. In: *to be published in Proc. Conference on Human System Interactions*, 2009

[DZL07] Deutsch, T. ; Zeilinger, H. ; Lang, R.: Simulation Results for the ARS-PA Model. In: *Proc. 5th IEEE International Conference on Industrial Informatics* Bd. 2, 2007, p. 995–1000

[DZLZ08] Deutsch, T. ; Zia, Tehseen ; Lang, R. ; Zeilinger, H.: A simulation platform for cognitive agents. In: *Proc. 6th IEEE International Conference on Industrial Informatics INDIN 2008*, 2008, p. 1086–1091

[Fen97] Fenichel, O.: *Psychoanalytische Neurosenlehre*. Bd. 1. Gießen : Psychosozial-Verlag, 1997

[Fre33] Freud, S.: *Gesammelte Werke; Neue Folge der Vorlesungen zur Einführung in die Psychoanalyse*. Bd. 15. Frankfurt am Mein : Fischer Verlag, 1933

[Fre40a] Freud, S.: *Gesammelte Werke*. Bd. 13. Frankfurt am Mein : Fischer Verlag, 1940

[Fre40b] Freud, S.: *Gesammelte Werke*. Bd. 1. Frankfurt am Mein : Fischer Verlag, 1940

[Fre40c] Freud, S.: *Gesammelte Werke; Triebe und Triebschicksale*. Bd. 10. Frankfurt am Mein : Fischer Verlag, 1940

[Fre84] Freud, A.: *Das Ich und die Abwehrmechanismen*. Frankfurt am Main : Fischer Taschenbuchverlag, 1984

[Gru07] Gruber, A.: *Neuro-Psychoanalytically Inspired Episodic Memory for Autonomous Agents*, Technische Universität Wien, Institut für Computertechnik, Diplomarbeit, 2007

[Köh93] Köhler, T.: *Das Werk Sigmund Freuds*. Bd. 2. Heidelberg : Roland Asanger Verlag, 1993

[Koh08] Kohlhauser, S.: *Requirement Analysis for a Psychoanalytically Inspired Agent Based Social System*, Technische Universität Wien, Institut für Computertechnik, Diplomarbeit, 2008

[LBP+07] Lang, R. ; Bruckner, D. ; Pratl, G. ; Velik, R. ; Deutsch, T.: Scenario Recognition in Modern Building Automation. In: *Proceedings of the 7th IFAC International Conference on Fieldbuses & Networks in Industrial & Embedded Systems (FeT 2007)* (2007), p. 305–312

[LZD+08] Lang, R. ; Zeilinger, H. ; Deutsch, T. ; Velik, R. ; Muller, B.: Perceptive learning - A psychoanalytical learning framework for autonomous agents. In: *Proc. Conference on Human System Interactions*, 2008, p. 639–644

[Min86] Minsky, M.: *The Society of Mind*. New York : Simon and Schuster, 1986

[Pal07] Palensky, B.: *Introducing Neuro-Psychoanalysis towards the Design of Cognitive and Affective Automation Systems*, Faculty of Electrical Engineering and Information Technology, Vienna University of Technology, Diss., 2007

[Pan98] Panksepp, J.: *Affective Neuroscience, the Foundations of Human and Animal Emotions*. Oxford University Press, Inc. 198 Madison Avenue, New York, 1998. – ISBN 0195096738

[PB07] Pfeifer, R. ; Bongard, J.: *How the body shapes the way we think*. MIT Press, 2007

[PB09] Palensky, Brigitte ; Barnard, Etienne: A Brief Overview of Artificial Intelligence Focusing on Computational Models of Emotions. In: Dietrich, D. (Hrsg.) ; Fodor, G. (Hrsg.) ; Zucker, G. (Hrsg.) ; Bruckner, D. (Hrsg.): *Simulating the Mind – A Technical Neuropsychoanalytical Approach*, Springer, Wien, 2009. – ISBN 9781412949644, p. 76 – 99

[PPC09] Palensky, P. ; Palensky, B. ; Clarici, A.: Cognitive and Affective Automation: Machines Using the Psychoanalytic Model of the Human Mind. In: Dietrich, Dietmar (Hrsg.) ; Fodor, Georg (Hrsg.) ; Zucker, Gerhard (Hrsg.) ; Bruckner, Dietmar (Hrsg.): *Simulating the Mind – A Technical Neuropsychoanalytical Approach*, Springer, Wien, 2009. – ISBN 9781412949644, p. 178 – 227

[Pra06] Pratl, G.: *Processing and Symbolization of Ambient Sensor Data*, Faculty of Electrical Engineering and Information Technology, Vienna University of Technology, Diss., 2006

[RLD+07] Roesener, C. ; Lang, R. ; Deutsch, T. ; Gruber, A. ; Palensky, B.: Action planning model for autonomous mobile robots. In: *Proc. 5th IEEE International Conference on Industrial Informatics* Bd. 2, 2007, p. 983–988

[Roe07] Roesener, Ch.: *Adaptive Behavior Arbitration for Mobile Service Robots in Building Automation*, Vienna University of Technology, Institute of Computer Technology, Diss., 2007

[SCS05] Kap. The Architectural Basis of Affective States and Processes In: Sloman, Aaron ; Chrisley, Ron ; Scheutz, Matthias: *Who Needs Emotions?: The Brain Meets the Robot*. Oxford University Press, Oxford, New York, 2005, p. 203–244

[Sol04] Solms, M.: *Das Gehirn und die innere Welt*. Düsseldorf und Zürich : Patmos Verlag GmbH & Co. KG Walter Verlag, 2004

[SSK97] Schuster, P. ; Springer-Kremser, M.: *Bausteine der Psychoanalyse*. Wien : WUV-Universitätsverlag, 1997

[Vel98] Velásquez, J.: Modeling emotion-based decision-making, 1998

[ZLM09] Zeilinger, H. ; Lang, R. ; Müller, B.: Bionic Inspired Information Representation for Autonomous Agents. In: *to be published in Proc. Conference on Human System Interactions*, 2009