Utrecht University

Master Artificial Intelligence

Master Thesis

The Logic of Mood

An Integration of Mood and Emotion in a Logical Model

Author
R.M. Bakker

Supervisors
Dr. M. M. Dastani
Prof. Dr. J-J. Ch. Meyer

May 16, 2018

Utrecht University

Abstract

Over the years we have seen an increasing amount of intelligent agent models with emotions. However, another affective state that is equally important, and which closely interacts with emotions, is often overlooked: *mood*. In this thesis, we aim to integrate the affect mood in models of agents with emotions. To gain a deeper understanding of the affective phenomena *emotion* and *mood*, we have explored theories of emotion and mood from different fields of study, mainly from psychology and artificial intelligence. The logical model of Steunebrink (2010) and the computational model ALMA are the foundations for a new structure of the process of emotion. This new structure allows for easy incorporation of mood. Upon closer inspection we find that there are discrepancies between the logical and the computational model. These models are both inspired by psychological theories. Therefore we investigate these underlying psychological theories of the models, and use the information from these theories to integrate mood with emotion. In the end we propose a new structure for the emotion elicitation process which combines factors from the original psychological theories and from the logical and computational models, and which eliminates discrepancies that occurred in the integration of mood and emotion.

Acknowledgements

I would like thank my supervisors, dr. Mehdi Dastani and prof. dr. John-Jules Meyer for their feedback, encouragement, and inspiration. I always walked away with a smile from our meetings, not only when the thesis was going well, but also when I mostly needed encouragement. Or when discussing whether sour was a nice taste. Their enthusiasm always inspired and motivated me. I would also like to thank Timo for hearing out my ideas, proofreading, giving excellent suggestions for solutions, and for encouraging me when things did not work out as expected. My thanks also goes to Minke for proofreading and suggesting rectifications. Furthermore, I thank my parents for their unconditional support and encouragement, without whom I would never have had this opportunity. Lastly I would like to thank everyone else who supported me in various other aspects while working on my thesis.

Contents

1 Introduction

Emotions have been a topic of interest since the ancient Greeks. Plato debated that humans must become solely rational by freeing themselves of emotion (Evans, 2002). This has long since been the most prominent view on emotions. In the second half of the 20th century, research on emotions became more popular, and the idea that emotions are a hindrance to rationality was soon questioned. Then several psychological and neurological experiments showed that emotions are an essential component of decision-making abilities (Bechara et al., 1994a). Some researchers took it even further, and stated that humans would not be able to display intelligent behavior without having emotions (Frijda, 1987; Ekman & Davidson, 1994; Bechara et al., 1994). Since then, the field of emotions and other affective processes (e.g. mood and stress) has been a large topic of interest, and it is striving towards a unified theory of emotion that could explain all its aspects. And it is not only the field of psychology anymore that showed interest in emotions, artificial intelligence soon followed.

Around the same time as when emotions started being a topic of scientific interest, a new area of expertise started to rise in the scientific world: artificial intelligence (AI). Philosophers and mathematicians had tried for centuries to capture parts of human thinking in logical and mathematical languages. A breakthrough came when the first digital computers came into existence in the 1940s. Ten years later, Turing (1950) asked the question that constituted the backbone of the field of artificial intelligence:

> 'I propose to consider the question, 'Can machines think?" (p.433)

This question is still relevant as of today. Several questions followed from it, like 'What is **a machine?**' Some philosophers stated that a human could be regarded as a machine as well. Another important question that was asked was 'What is **thinking**?' These questions led to philosophical debates about human intelligence and ontology as regards to machines. In this thesis, we do not pursue the ethics and philosophical questions of artificial intelligence, but it is important to keep in mind that notions such as intelligence and consciousness are not self-evident.

As of today, intelligent agents are becoming more advanced and human-like by the year. In this thesis, *intelligent agent* or shortly *agent* is used to denote a software entity that performs actions in a relatively independent way. The first computational models with emotions sprouted around the 1990s, and there have been many additions since. Their aim differs from making their model more human-like to improving decision making in agents. Meyer (2004) introduced the first logical model of emotion, starting with a couple of very basic emotions, and using the KARO framework to express them in logic. Most of these agent models treat emotions and leave other affective processes, like mood, in the dark. This thesis aims to integrate emotion and mood in one logical model. We[1] want this integration to be based on multiple disciplines and not to be model-specific, such that it can

[1] I will use 'we' throughout this thesis as *pluralis modestiae*

be used by other models. This is done by examining how emotions and moods are expressed in psychological, computational, and logical models and theories, by comparing these models and theories, and in the end by combining two different models based on two different theories.

1.1 Motivation

Until recently, emotions were seen as a hindrance to the human mind, they clouded the mind and prevented humans from thinking rationally. After the 1950s this view changed, although the intuition of many people still remains that emotions are the opposite of a rational mind. In Bechara et al. (1994b) neurological evidence indicated that emotions are essential to human intelligence; not having enough emotions or emotions not functioning correctly can have great impact on the behavior and actions one makes. Bechara et al. (1994b) studied patients with damage on the prefrontalamygdala circuit. This place in the brain is responsible for memory and for emotion. Thus, the patients with these conditions had impaired emotional capabilities. The result was that these patients were highly impaired in decision making. A relatively simple task such as planning a dentist appointment became almost impossible. Therefore it can be concluded that emotions play a big role in decision making in humans. Some take it even further: Salovey and Mayer (1990) and later Goleman (1995) state that emotions are a crucial part to intelligence in humans. Emotional processes such as the regulation of stress, motivation, and 'the ability to empathize and to hope' are such important aspects that they would predict success in life better than traditional IQ tests (Goleman, 1995).

Since emotions are apparently so important for humans, one could reason that therefore they should also be taken into account in AI. The philosophical idea of *strong AI* aims at creating intelligent agents that are as human-like as possible. In this view adding emotions would make the agent more human-like, regardless of the benefits of the emotions. The benefits would of course also be a reason for adding emotions to AI. By modelling emotions in AI, a better theoretical understanding of emotions can be achieved, and at the same time the architecture of intelligent agents would be enriched (Reisenzein et al., 2013). As mentioned above, emotions aid decision making in humans, therefore it could also provide an alternative for rational decision theory in intelligent agents (Dastani et al., 2014). Another argument for adding emotions to AI is that emotions in agents improve the believability of an agent, and they improve the interaction between humans and agents (Velásquez, 1996; Mehrabian, 1996a).

Emotion is not the only affective process that plays an important role in human cognition, other affective processes are for instance stress, mood, and personality. They are mostly differentiated by their duration, stress being the shortest and personality the longest. They interact closely and have great influence on each other (Lazarus, 1991). Stress, mood, and personality are concepts from psychology that have not been adopted yet by AI in the way emotions have. Especially mood and emotions interact in important ways and therefore should be considered in an affective agent model (Velásquez, 1996). Mood can influence emotions and

other cognitive processes in many ways, for instance, a positive or negative mood can result in forgetting, or remembering more than in a neutral mood (Clore and Huntsinger, 2007; Bäuml and Kuhbandner, 2007). Mood also influences emotions, when a human is in a negative mood, negative emotions are easier triggered, and when it is in a positive mood, positive emotions are easier triggered (Steunebrink, 2010).

Shortly summarized, intelligent agent models would benefit from having emotions and mood because it makes them more like humans, it can aid their decision making, it will make them more believable and therefore also improve the interaction with humans, it will give a better theoretical understanding of emotions, and it will enrich the models.

1.2 Research Questions

The main research question of this thesis will be:

- How can mood be integrated in models of agents with emotions?

To answer this question, we can split it up into multiple smaller subquestions:

1. What are the affective phenomena *mood* and *emotion* exactly, how do they differ, and what is their relation?
2. How are *emotion* and *mood* formalized in agent models?
3. How can *emotion* and *mood* be integrated?

1.3 Methodology

To answer the research questions, first a literature study will be done in chapter 2. We will analyze theories and models from psychology, computer science, and artificial intelligence to get a complete view of the different uses and definitions of emotion and mood. At the end of chapter 2 we will answer the first two subquestions. In chapter 3 we try to integrate mood and emotion in one model, using ideas and theories from two different models of emotion and mood. We first discuss why mood and emotion should be integrated in one model, and why existing models do not allow this. Then we explain what is needed for mood and emotion to exist in one logical model, and propose a new structure of the process of emotion that accommodates these needs. After this we will zoom in on the emotion elicitation process, and make adjustments to this process such that mood can be added. In chapter 4 we will analyze and improve the findings from chapter 3. We will look at the original psychological theories that underlay the computational and logical model used in chapter 3, and see how they differ. Lastly, we will end with chapter 5, which summarizes this thesis and answers the research question. We will also reflect on our work and make suggestions for future work.

2 Related Research

In this chapter we are going to give an overview of the field of emotion and mood. We will also explain several logical and computational models that are based on psychological theories of emotion. First we will try to find an answer to our first sub question by comparing literature on what emotion and mood exactly are, and how they relate in sections 2.1.1 and 2.1.2. Then we will give a more detailed account on the psychological theories and computational models that are most relevant to answering our main question. This should also enable us to answer the first and the second subquestions. We will shortly summarize the chapter and answer the sub questions in 2.4.

2.1 Emotion, mood, and the difference between them

Emotion and closely related topics like stress, mood and coping have been studied in different fields. Sometimes these relate to each other, but more often the fields of study all have their own theories, models, and definitions that do not align with the ones of other fields. The fields of study that are relevant to this thesis are psychology, philosophy, artificial intelligence and computer science. In those fields, there are again different tendencies on emotion. So, to narrow it down, we make a rough distinction between these theories from different fields. This should help to give some structure and insight to the vast pool of theories and models on the topic. A schematic view of different theories is given in figure 1. There have been many psychological theories on emotion. The ones that are still relevant as of today and are mentioned in this chapter are shared under the header 'Psychological theories'. Since the proposal of a logic of emotions for intelligent agents by Meyer in 2004, several logical models have been proposed. They are gathered under 'Logical Models' in figure 1. Partly based on psychological theories and partly on logic, computational models arose around ten years ago. They are displayed under 'Computational Models'.[2]

In this thesis we first want to get a working definition of mood and emotion. To accomplish this we investigate several theories from different fields of study in the next sections. After this analysis on a definition, one will be chosen according to our purposes.

2.1.1 What is emotion?

Many different answers can be found to the question 'What is emotion?' What answer is found depends highly on where one searches for an answer, since a precise definition of emotion can only be given in the context of an emotion theory (Reisenzein et al., 2013). Nevertheless, we want to give a short overview of definitions used to give the reader some idea of what motivates the use of a definition

[2] **Psychological theories:** Lazarus (1966, 1993); Folkman and Lazarus (1988); Frijda (1987a,b); Frijda et al. (1989); Ortony et al. (1988); Clore and Huntsinger (2007) **Computational models:** Gratch and Marsella (2004); Gebhard (2005) **Logical models:** Meyer (2004, 2006); Adam (2007); Steunebrink et al. (2007, 2009); Steunebrink (2010)

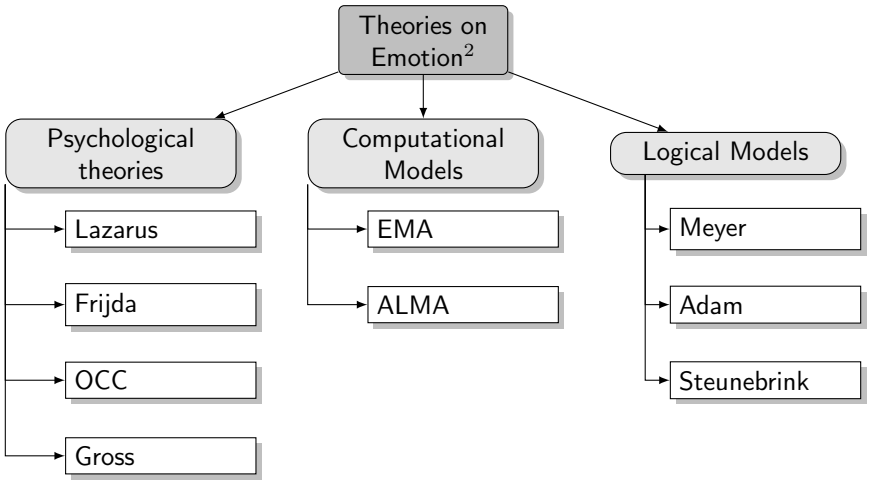

Figure 1: A schematical view on different theories on emotion

in this thesis. Since the field of artificial intelligence is one of many different disciplines, we will study literature from various fields to find a definition of emotion. We will begin our search with definitions of emotion from psychology, since this is the field where the modern emotion theories originated. Then we will analyze some of the definitions used in computational models of emotion, and lastly the definitions used in logical models will be studied.

Psychological theories Lazarus (1966) was one of the first psychologists to recognize the chaos in the fields of emotion and stress. What are the phenomena exactly and how do they differ? This was a question often untreated and therefore many scientists used their own idea of emotion, stress, and related affects. In one of his more recent studies (Lazarus, 2006) it is mentioned that the distinction is still not that clear and for a reason: emotion, stress and related mechanisms like coping are so closely related that a hard distinction cannot be made. The relationship between emotion and another important affect, mood, is also not always clear. This relationship will be treated in section 2.1.2. For now let us focus on a definition of emotion on its own.

Many psychologists followed the example of Frijda (1987b) in defining a definition of emotion. He stated that there cannot exist a definition of emotion without a complete theory, and follows by designing that theory. In line with his theory on emotion he concludes that 'Emotions are defined as changes in action readiness. ...' (p. 466). This is a more physiological explanation of emotions. Around the same time Folkman and Lazarus (1988) used a definition that is alike but a little more specific:

"We have defined emotions as complex, organized psychophysiological

reactions consisting of cognitive appraisals, actions impulses, and patterned somatic reactions." (p.310)

Here emotions are not only changes in action readiness, or action impulses as Folkman and Lazarus (1988) call it, but also cognitive appraisals and somatic reactions that work as an unit. This is not only a physiological view of emotion but also psychological. In section 2.2.1 we will discuss cognitive appraisal further since it is an important theory on emotion elicitation that is also used in computational models. For now it suffices to say that the definitions on emotions from both Frijda (1987b) and Folkman and Lazarus (1988) are based on physiological and psychological reactions, where cognitive appraisal is a central component to this reaction. But why is this the case? This question was explored by Ortony et al. (1988).

A model of emotion elicitation was introduced by Ortony et al. (1988), and based on this model the following definition of emotion was used:

> 'Our working characterization views emotions as valenced reactions to events, agents, or objects, with their particular nature being determined by the way in which the eliciting situation is construed.' (p. 13)

In short, emotions are positive or negative cognitive reactions to either events (e.g. **fire breaks out** and one feels fear), agents (e.g. A father is proud of **his daughter**), or objects (e.g. Someone loves **his new watch**). This definition is made more specific in their model, which elaborates on how reactions are positive or negative and also adds other levels of distinction below events, agents, and objects. The difference between this definition of emotion and the ones proposed by Frijda (1987b) and Folkman and Lazarus (1988) is that this definition tries to make both the process of emotion elicitation and the emotions themselves more transparent. One important reason for this is that the goal of Ortony et al. (1988) was to find a cognitive structure of emotion, whereas Frijda (1987b) aims to fully understand emotions as a whole and its effects, and Folkman and Lazarus (1988) reasons from multiple aspects of emotion like social influences.

One concept that the above definitions of emotion have in common, is that they all mention that emotion is some reaction toward something in the environment of the subject. The model of Ortony et al. (1988) is most specific, but also covers the smallest part of of emotion theory in general. However, since it is most specific, and furthermore they have also tried to lay a foundation for a computationally tractable model of emotion, this definition is often used to formalize emotion in both logical and computational models.

Computational models In computational models, *emotion* is often a slightly modified concept taken from a psychological theory of emotion. For example, the computational model EMA (Gratch and Marsella, 2004) heavily relies on appraisal theory (see section 2.2.1) in general and more specifically uses ideas on emotion from Smith et al. (1990), Frijda (1987a), and Ortony et al. (1988). In Gratch

and Marsella (2014), an overview of computational models of emotion and their corresponding psychological theories is given. This overview is shown in figure 2.

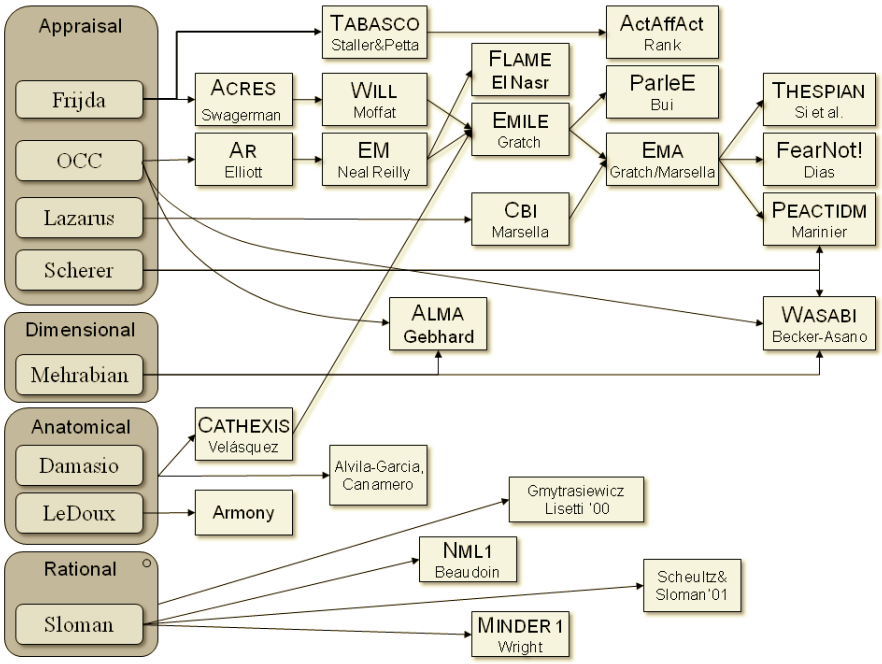

Figure 2: An overview of computational models with on the left hand the psychological theory of emotion they are based on. Adapted from Gratch & Marsella, 2014, p.6

As can be seen in figure 2, many computational models are influenced by multiple psychological theories. This in turn also influences the definition of emotion that they took on in their model. Often a precise definition is never mentioned, rather the theory behind it. The theory of emotion and a rough working definition is more important than a precise definition for conducting meaningful research, according to Reisenzein et al. (2013). A working definition of emotion can minimally be provided by a list of examples, like mental states that are denoted as 'anger', or 'joy'. Uncontroversial characteristics of emotion can also be added. There are some features that are widely accepted, like the feature that emotional states are reactions to some perception, or a thought. Another feature is that emotions can be directed at some object or an event (Reisenzein et al., 2013). In the work of Reisenzein et al. (2013), their working definition of emotion is extended with the assumption that emotions are mental states that cause new mental states and behavior.

In another computational model of emotion, Cathexis, a comparable view is

adopted on defining emotion (Velásquez, 1996). They want to give an idea as to what emotion is rather than to give a precise definition. However, several characteristics are pointed out that are widely accepted. The first feature of emotion mentioned is more or less the same as the first characteristic Reisenzein et al. (2013) mentioned: emotion corresponds to an elicitor, or activator. This is another way of saying that an emotional state is a reaction to an outside or inside stimulus. A second feature that is mentioned is different: Velásquez (1996) states that emotions include an expressive component to help define it (e.g. smiling). This second characteristic is not one that all researchers in this field agree on, some see it as an essential attribute, others mention it but do not treat it in depth. This again depends on the specific model and its purpose. For instance, the expressive component of emotion is a central subject in the Cathexis model, but not in the analysis of Reisenzein et al. (2013).

Sometimes the purpose of a model is clear, for example in ALMA (A Layered Model of Affect) (Gebhard, 2005). This model was part of project that developed interactive virtual humans with which a conversation could be held. In this model, emotion is described as:

> 'Emotions reflect short-term affect, which is usually bound to a specific event, action or object, which is the cause of this emotion. After its elicitation emotions usually decay and disappear of the individual's focus.' (p. 3)

The definition of emotion used here is precise and practical as to make it suitable for implementation. Note that this definition has similarities to the first feature described by Reisenzein et al. (2013). It is however formulated in a way more fitting to the model of Gebhard (2005).

Even though computational models of emotion are based on different psychological theories, we can see similarities in what they deem to be emotions. A recurring feature of emotion is its directedness to some event, thought, or object. As you might remember, this was also a common characteristic of psychological theories. Another attribute of emotion that returns more often in computational models is the duration. This aspect will be treated more thoroughly in section 2.1.2.

Logical models In this thesis, we aim to integrate mood and emotions in a logical model. Hence we will also shortly treat definitions of emotion in existing logical models, even though they are very similar to computational models.

Meyer (2004) was the first to propose a formal logic for emotions and related affects. Here emotions are viewed as a structuring mechanism, a tool to help intelligent agents organize and prioritize actions. Therefore they do not necessarily exactly match the way emotions are experienced by humans. They only initiate expected behavior in a state of some emotion (Meyer, 2004; Picard, 1997). The question whether these emotions are like emotions in humans, and if they really experience them like humans do, is not relevant for the design of intelligent agents (Meyer, 2004). However, the emotional states used do naturally relate on some level to emotions in humans.

Ideas from the psychological theory of Ortony et al. (1988)(the OCC model) are used in Meyer (2004), but it mainly follows Oatley and Jenkins (1996). This is also the case for the model of Dastani et al. (2006). The formalizations of Steunebrink (2010), Adam et al. (2009) and others (Steunebrink et al., 2007; Adam, 2007), are completely based on the psychological theory of Ortony et al. (1988). These formalizations try to model and formalize human emotions and their quantitative properties, and therefore heavily rely on the OCC model (section 2.2.3). The notion of emotion is therefore more like the notion as it is in the OCC model. There is, however, little attention as to a precise definition of emotion; the logical models, like the computational models, focus more on the function of emotion than on what they are exactly.

Steunebrink (2010) does shortly treat the question as to what emotion is exactly, for this he relies on both the OCC model and the theory of emotion regulation from Gross (2013). This last theory adds another component to emotion: its duration. The duration of an emotion can mainly help to distinguish it from other affective phenomena, like *mood*. In the next section we will explore the further facets of mood, and how it differs from emotion.

2.1.2 What is mood?

Now that we have given some idea to the reader about the concept of emotion, we will analyze another affective phenomenon for which emotion is very important: *mood*. As another notion used frequently in daily life, most people have a general idea as to what *mood* is. But how would one define it precisely? There is no consensus about this, just like it is with *emotion*. There are again different views in different fields of study on the topic, we will shortly discuss them below.

Psychological theories One view on mood and emotion that is shared by many psychologists, is that emotion is directed. That is, the emotion arises due to some event, action or object. Mood on the other hand is not. A definition that is still used by some psychologists as of today was provided by Nowlis (1970). He stated that mood is an 'effect upon the person of his own configuration of activity', which is another way of saying that mood does not emerge from objects, actions or events but from someone's mind. Furthermore the mood can influence other processes such as the level of concentration, social orientation, and positively or negatively influence general appraisal (Nowliss 1965, p. 353) .

This view is one that is shared among many psychologists (e.g. Frijda (1987b); Gross (2013); Clore et al. (2001)). There are however psychologists that have different ideas about mood and emotion in general. Mehrabian (1980) proposed to introduce 'basic dimensions' for the social sciences. The study of affects and emotions is in his view handicapped because of behaviorist bias. In contrast with the social sciences, natural sciences use a set of basic dimensions, such as mass, time, or length. Mood is, in his view, the average emotional state across 'a representative sample of life situations' (Mehrabian, 1996b). Otherwise he agrees with the idea that mood is not directed towards some event.

Computational and logical models Affective computing is a relative new field of study, and most computational and logical models are therefore focused on the affect that is best studied in psychology: emotion. Some computational models do not treat mood at all, others do to more or less extend. The ones that do usually are heavily based on psychological theories, however, they focus more on the function of mood than on the precise concept. One example of such a model is Cathexis (Velásquez, 1996). In Cathexis, it is mentioned that moods last longer than emotions, and are continuous of nature. Mood is here explained by its functions. For example, in Cathexis, mood lowers the activation threshold for certain emotions, that is, some emotions will arise more easily when a certain mood is present. They also treat the cause of mood. How does a mood establish itself? One explanation is changes in one's biochemical state. Another that is mentioned by Velásquez (1996) is the occurrence of dense emotional episodes. The design of emotion is done according to this perspective: arousal of emotion might also activate a mood state that in turn can influence the next arousal of emotion. Most computational models that incorporate mood have views similar to the ones in Cathexis, but they usually don't treat the exact cause of mood. We will describe other examples in the next section. In logical models, moods are sparsely used. Steunebrink (2010) uses mood in a very simplified version of the Cathexis model as we will see in the next section.

There is general agreement in all of the fields that emotions are shorter in duration than moods. Oatley and Jenkins (1996) noted that duration is a distinguishing factor in all affects. They illustrated this with figure 3.

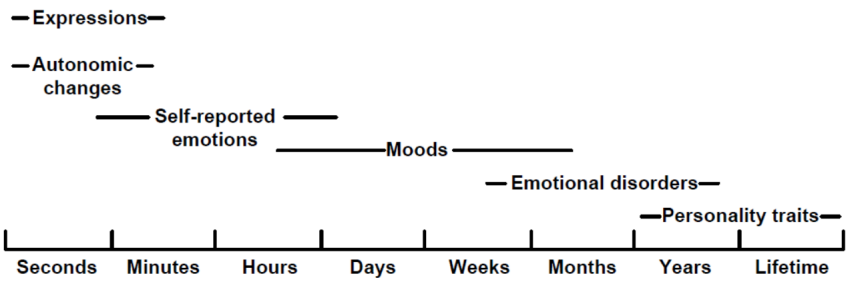

Figure 3: A spectrum of affective phenomena and their typical duration. Figure taken from Oatley and Jenkins (1996), p. 124

According to Oatley and Jenkins (1996), emotions have a duration of minutes or hours, while moods can linger for months. The duration is not always agreed on, for example Mehrabian (1980) thinks that moods are closer to what Oatley and Jenkins (1996) call personality traits, while McCrae and John (1992) place moods closer to emotions in their duration.

2.2 Psychological theories

Now that we have given a global idea of what mood and emotion is, we are going to analyze some concepts and theories that we will need in the coming sections. We will begin by briefly introducing appraisal theory, one of the most influential theories that underlay a lot of other theories on emotion in psychology. Then we will give a short outline on the PAD theory and on the OCC model. The first is an influential theory on basic dimensions for affective processes, the second is a model which captures the essence of emotions in a modular way and is picked up by many scientists from other fields, like AI.

2.2.1 Appraisal theory

There are many forms of appraisal theories, but they have several things in common. The first is the assumption that emotions arise by daily life situations. A human encounters some stimulus, then it interprets the situation, then this interpretation is processed by the brain, and an emotion arises (Smith et al., 1990). The interpretation of the situation is called *appraisal*. A highly abstracted and simplified form of the appraisal process is given in figure 4. An example of the appraisal process would be the following: You run into a friend that you have not seen for a long time (situation). Then you remember why you have not seen him in two years: he borrowed quite some money and did not pay you back. However, in the mean time you won the lottery, and after all, what is money compared to a friendship? This would be a chance to restore the friendship (1). Or, in the mean time you struggled to make ends meet, and you cannot believe he has not contacted you during these troubled times (2). This weighing of the situation towards your own mental state is appraisal. The emotion that might arise in case 1 could be joy or relief. Joy about seeing a friend again, or relief that you can put this fight between you in the past. In case 2, you would probably feel anger or disappointment. Appraisal is essential to what emotions will arise in this situation (Moors et al., 2013).

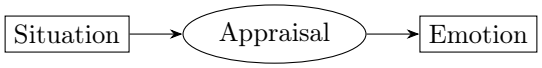

Figure 4: A simplified graphic representation on the appraisal process

There are many different theories based on appraisal theory, some of the most influential being the OCC model (Ortony et al., 1988), Frijda's theory (Frijda, 1987b), Lazarus' theory (Lazarus, 1991), and the work of Scherer (Scherer et al., 1984). Frijda's theory and the OCC model are the most influential for computational models. We saw this in figure 2. The main difference between those two is the focus of their studies: The OCC model tries to provide a structure as to how emotions work (see section 2.2.3), Frijda's theory aims to fully understand emotion, that is, its definition, its function, relations with other cognitive processes,

etc. There are many other differences between these theories, exploring all of them would go beyond the scope of this thesis. For now the most important distinction between theories like the OCC model, and others like Frijda's theory or Lazarus' is their goal: providing a clear structure vs. giving a complete and truthful account of emotions. Appraisal theory together with having a clear structure, like the OCC model has, lends itself perfectly for logical and computational models. However, appraisal theory and its subtheories are not the only way of thinking about emotions, as we will see in the next section.

2.2.2 The PAD theory

The three-factor theory of emotions, the three factors being *pleasure* (P), *arousal* (A), and *dominance* (D), aims to provide a basic set of dimensions for the 'softer' fields of psychology (personality, social, developmental, environmental) (Mehrabian and Russell, 1974). In this thesis we shall refer to it as 'the PAD theory'. In psychology, the study of human emotions was long hindered by historical bias. As mentioned earlier, emotions were seen as a hindrance to ratio, and therefore should be limited. Behaviorism also did not help with the study of emotions. Therefore the field was quite scattered with terms and theories. The PAD theory tries to change this by proposing that psychology should have basic dimensions, like the physical sciences have length, time, and mass (Mehrabian, 1980). This theory can therefore be called a *dimensional* theory and can be separated from theories with other basic assumptions, like appraisal theories that depend on the concept of appraisal.

The basic dimensions that are proposed are *pleasure*, *arousal*, and *dominance*. Pleasure, and its counterpart displeasure, can be recognized by positive versus negative facial expressions. It can be seen as a continuum ranging from extreme happiness to extreme unhappiness, with adjectives as happy vs. unhappy, and pleased vs. annoyed (Bakker et al., 2014). The dimension arousal/nonarousal is a combination of activity and alertness. A high state of arousal would for example be a tennis player, who moves across the field (activity) and thinks of a strategy (alertness) (Mehrabian, 1980). A low state of arousal is for example someone who is sleeping. Adjectives that best describe arousal are stimulated vs. relaxed, and excited vs. calm. Someone's dominance is based on the extent to which the person feels free to act. Dominance can be influenced by settings which limit the persons actions and control. The counterpart of dominance is submissiveness, where a person feels very restricted in his actions and not in control. Together, these basic dimensions can account for all emotional reactions.

There were several experimental studies done to determine the emotional descriptors, the specific values of the PAD dimensions, and to test the independence of the dimensions. In one study, 200 subjects were given a description of a situation. Then they were asked how they would feel in that situation, and to describe his feelings according to three scales: pleasure, arousal, and dominance (table 1). Finally subjects would fill the adjectives they found suitable for the situation in a 255-item checklist. In a second study, 300 subjects were asked to define emotion

denoting terms directly, using the same scales again. This way the relationship between the exact term and the PAD values could be assessed (Russell and Mehrabian, 1977).

Pleasure										
Happy	_	_	_	_	_	_	_	_	_	Unhappy
Pleased	_	_	_	_	_	_	_	_	_	Annoyed
Satisfied	_	_	_	_	_	_	_	_	_	Unsatisfied
Contended	_	_	_	_	_	_	_	_	_	Melancholic
Hopeful	_	_	_	_	_	_	_	_	_	Despairing
Relaxed	_	_	_	_	_	_	_	_	_	Bored
Arousal										
Stimulated	_	_	_	_	_	_	_	_	_	Relaxed
Excited	_	_	_	_	_	_	_	_	_	Calm
Frenzied	_	_	_	_	_	_	_	_	_	Sluggish
Jittery	_	_	_	_	_	_	_	_	_	Dull
Wide awake	_	_	_	_	_	_	_	_	_	Sleepy
Aroused	_	_	_	_	_	_	_	_	_	Unaroused
Dominance										
Controlling	_	_	_	_	_	_	_	_	_	Controlled
Influential	_	_	_	_	_	_	_	_	_	Influenced
In control	_	_	_	_	_	_	_	_	_	Cared for
Important	_	_	_	_	_	_	_	_	_	Awed
Dominant	_	_	_	_	_	_	_	_	_	Submissive
Autonomous	_	_	_	_	_	_	_	_	_	Guided

Table 1: Measure of emotional responce according to PAD scales (Mehrabian and Russell, 1974)

Pleasure and arousal scales intercorrelated 0.03, pleasure and dominance 0.40, and arousal and dominance 0.15. These intercorrelations are not incompatible with their assumed independence, because nonzero linear relationships may be seen among any particular sample of emotional states. The correlations were also sufficiently low to permit regression analyses (Russell and Mehrabian, 1977). In the second study, scores were computed in terms of pleasure-displeasure, arousal-nonarousal, and dominance-submissiveness according to the rating on the scales they were given. Table 2 shows an example of the scores, full scores can be found in Appendix C. The scales for the P, A, and D values ranged from -1 to +1, with a neutral value 0. For each term the number of subjects who provided the data is given (N), the mean for each dimension and the standard deviation for each dimension. The reliability of the mean scores were estimated and yielded reliability coefficients of 0.97 for pleasure, 0.89 for arousal, and 0.87 for dominance.

Russell and Mehrabian (1977) argue that, despite some discrepancies in the exact terms, these studies provide enough evidence to show that the three inde-

pendent dimensions are both necessary and sufficient conditions for an adequate description of emotions. However, their study also shows that the dimensions are not completely independent, just enough so that a difference can be made between them. The terms are also not always in accordance with expectancy or previous studies, for instance *friendliness* and *affectionate* had slightly negative arousal values in their first study, and highly positive arousal values in the second study. Other discrepancies occurred with *aggression*, *anxious*, and *distress*.

		Pleasure		*Arousal*		*Dominance*	
Term	N	Mean	SD	Mean	SD	Mean	SD
1. Bold	27	.44	.32	.61	.24	.66	.30
2. Useful	27	.70	.20	.44	.28	.47	.40
3. Mighty	27	.48	.37	.51	.28	.69	.31
4. Kind	27	.73	.22	.19	.32	.57	.27
5. Self-Satisfied	27	.86	.10	.20	.40	.62	.31
6. Admired	29	.81	.21	.44	.30	.51	.34
7. Proud	29	.77	.21	.38	.34	.65	.33
8. Interested	29	.64	.20	.51	.21	.17	.40
9. Arrogant	29	.00	.51	.34	.44	.48	.34
10. Inspired	29	.71	.30	.63	.21	.34	.55

Table 2: Part of the terms denoting emotions in terms of Pleasure, Arousal, and Dominance (Russell and Mehrabian, 1977).

The PAD theory is also applicable to other affective processes, such as *personality* and *mood*. Mehrabian (1978, 1980) use the term temperament to account for both personality and mood. They only make the distinction in *state* and *trait*, where state stands for a short term affect that is present for a select amount of time, and trait for a long term affect. Applying the PAD theory to temperament results in three measures of personality: The trait pleasure, which correlates with social desirability, achieving tendency, extroversion, and affiliative tendency, the trait arousal, which correlates to anxiety, neuroticism, and aggression, and the trait dominance, which correlates to autonomy, achieving tendency, and aggression. These three traits together result in a three dimensional space, from which the octants as a whole can be seen as temperaments. Mehrabian and O'reilly (1980) identified these octants through regression analysis as *exuberant* (pleasant, aroused, dominant) , *anxious* (unpleasant, aroused, submissive, *dependent* (pleasant, unaroused, submissive), *aggressive* (unpleasant, aroused, dominant), *depressed* (unpleasant, unaroused, submissive), *disdainful* (unpleasant, unaroused, dominant), *docile* (pleasant, unaroused, submissive), and *relaxed* (pleasant, unaroused, dominant). These traits come into existence by 'an average of a person's emotional states across a representative variety of life situations' (Mehrabian, 1996b). This wording leaves it open to the interpretation of others as what is a representative variety.

2.2.3 The OCC model

The OCC model, named after its authors, was introduced in 1988 (Ortony et al., 1988). It tries to give an account of how appraisals are made by providing a structure of emotion eliciting conditions. A central question in this work is the question of what distinguishes one emotion from another. This question is answered by providing a structure where different kinds of cognitive phenomena differentiate between emotions. The authors do not aim to give a complete and correct structure of emotion elicitation, but rather to give an idea as how to sort emotions according to their eliciting conditions, and do that in such a way that this structure is suitable for incorporating in AI theories.

In the OCC model, emotions are structured by their eliciting conditions. Emotions that have similar eliciting conditions are placed in the same group. For instance, the group 'well-being' contains the emotions *joy* and *distress*, and is separated from the group 'prospect based' by having the eliciting condition 'prospects irrelevant'. A visual view of this structure is given in figure 5. In interpreting the figure, labels in upper case represent structural elements, while labels in lower case represent emotional states. This structure starts with the separation of a valenced reaction into three aspects, *events*, *agents*, and *objects*. A valenced reaction is a reaction that is either positive or negative. *Events* are things that happen, *agents* are entities that can cause or contribute to events, *objects* are objects viewed qua objects (Ortony et al., 1988). Then one can be *pleased* or *displeased* about a consequence of an event, *approving* or *disapproving* about an action of an agent, or one can *like* or *dislike* an aspect of an object. When a consequence of an event is focused on other(s), the group 'fortunes-of-other' is applicable. What emotion is then elicited depends on whether the event is desirable or undesirable for the other. When a consequence of an event is focused on oneself, the prospects of the event matter. When the prospects are relevant, the emotions *hope* and *fear* are elicited. When this hope or fear is in turn confirmed, *satisfaction* and *fears-confirmed* are respectively elicited. When hope or fear is disconfirmed, *relief* and *disappointment* are elicited. When consequences for oneself have irrelevant prospects, *joy* or *distress* is elicited. The rest of the figure works in a similar way.

The total amount of emotions in this model adds up to 22, excluding the emotional states directly following the *events*, *agents*, *objects* distinction. Some of these emotions are more complex than others, however, Ortony et al. (1988) want to avoid talking about the simpler ones as 'basic emotions'. They prefer to think in terms of levels of differentiation, for instance the most 'basic' level of differentiation would be the two affective reactions (positive and negative). One level below that we can make another distinction in three basic classes of emotions: a positive or negative reaction to an event (pleased vs. displeased, a positive or negative reaction to an agent (approving vs. disapproving), and a positive or negative reaction to objects (liking vs. disliking). For all the 22 emotion types, specifications are given. An example is given below in figure 6.

An emotion specification has a type identification (i), which serves as a label for the type of emotion. It also has a type specification (ii), which contain the necessary

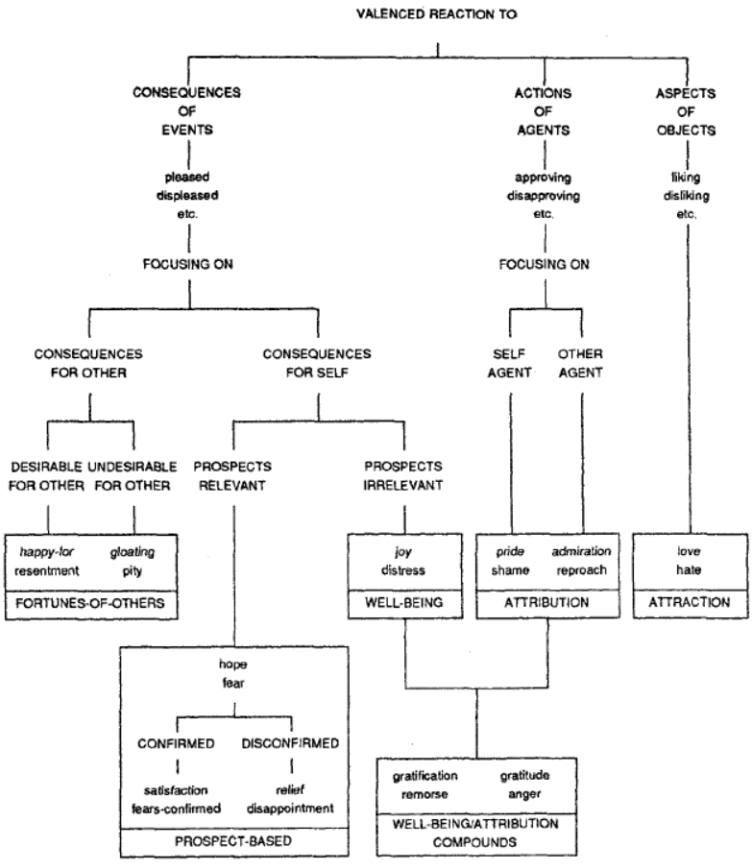

Figure 5: Global structure of emotion types, copied from the book (Ortony et al., 1988, p.19)

conditions for the experience of the emotion. Tokens (iii) are a list of words or phrases that describe the family of emotions of this type, they share the same type specification. (iv) Denotes variables affecting intensity, that is, the factors that influence the intensity of the emotion. In this case, the desirability of the event determines the intensity of joy. Lastly, the emotion specification concludes with a prototypical example (v). The other emotion specifications from the OCC model can be found in Appendix A.

Intensity is determined by so called 'global variables' and 'local variables'. Global variables are variables that influence all emotions, for instance *unexpectedness* and *arousal*. Local variables influence only certain emotions, e.g. *desirability*, *likelihood*, or *effort*. Desirability is a variable only relevant to consequences of events, an not to aspects of objects. likelihood influences hope and fear, the intensity depends on how likely the person thinks it is the feared or hoped situation

(i)Joy Emotions
(ii)TYPE SPECIFICATION: (pleased about) a desirable event
(iii)TOKENS: contented, cheerful, delighted, exstatic, elated, euphoric, feeling good, glad, happy, joyful, jubilant, pleasantly surprised, pleased, etc.
(iv)VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event is desirable
(v)EXAMPLE: The man was pleased when he realized he was to get a small inheritance from an unknown distant relative.

Figure 6: Emotion specification of *joy*, copied from (Ortony et al., 1988, p.87)

will arise. In total there are four global variables, and twelve local variables, to determine the intensity of all the emotion types.

The OCC model provides some structure and order in different emotions and what distinguishes them. It also gives an idea of factors that influence the intensity of an emotion. Whether this is a correct and complete approach remains uncertain, but that was not the aim of this model. There are several issues left open in this theory, mainly the experience of emotions, unconscious emotions, and coping. These issues also highly depend on personal differences, which is another topic in itself. However, this model succeeds in making emotions less cluttered, and in providing a basis for AI applications, as we will see in the next section.

2.3 Computational and Logical models

Since the 1990's there has been an increasingly amount of computational models and logical models. They are usually based on psychological theories, and have become more advanced over the years. The goals differ, some models aim to display emotions, some to recognize them, others to structure them (mostly the logical models), and combinations of these aims. There are too many models to treat all of them, so we only treat those that are influential for the field and of importance for the coming chapters. For comparisons and an overview of other models see for example Reisenzein et al. (2013); Marsella et al. (2010).

2.3.1 Computational models

A computational model based on the OCC model, PAD theory, and the five factor model of personality, is ALMA, 'A Layered Model of Affect' (Gebhard, 2005). The model is part of a project that develops interactive virtual characters for human-like conversation. Therefore the aim of the model is to help control behavioral aspects of the virtual character through emotions. ALMA differentiates three kinds of affects: emotions (short-term), moods (medium-term), and personality (long-term). The emotions are based on the OCC model, and implemented in java through an *emotion engine* (Gebhard et al., 2003a, 2004). Mood is based on the PAD theory, and is calculated with a *mood engine*. For personality, another

psychological theory is used, namely the five factor model (McCrae and John, 1992). The process of the emotion arising differs a little from appraisal theory, mainly because of the way the application is build. ALMA can be split into two phases, a preparation phase and a runtime phase. In the preparation phase personalities and appraisal rules are specified using *dialog act tags* and *appraisal tags*. These tags are abbreviations of *emotion eliciting conditions* (EECs). In the runtime phase, the tags are transformed into EECs, which in turn serve as input for the emotion engine. The emotion engine takes as input the EECs, a personality of the agent, and the current mood of the agent. The emotion engine then calculates the emotion(s), and the agent will experience the emotion(s). Mood is calculated in a similar way. It takes the personality and emotions of an agent as input, calculates the average PAD value of the emotions together with the personality, which lead to a current mood octant, and an exact value of mood to illustrate the intensity of the mood. A graphical representation of this process is given in figure 7.

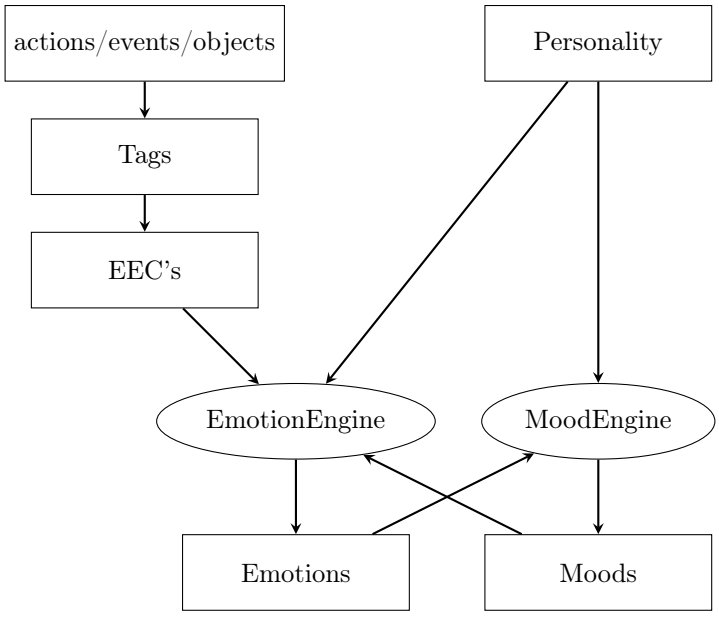

Figure 7: The process of emotion from ALMA, where square boxes are data and rounded boxes processes

The same emotions as the OCC model specifies are used, together with their eliciting conditions. It is however 'flattened out'; the eliciting conditions are not used in a modular way as Ortony et al. (1988) intended. The moods used are similar to those in the PAD theory, accept *aggresive* is changed to *hostile*, and *depressed* to *bored*. No motivation is given for this change. Since emotions serve as input for the calculation of mood, they have been given PAD values. The same

goes for the personality.

Another computational model with emotion is the EMA model. The EMA model aims to develop a general computational model of the mechanisms underlying human emotion. This helps both the understanding of human behavior in general, as well as advances the field of emotion in artificial intelligence (Gratch and Marsella, 2004; Marsella and Gratch, 2006). The main goal of the EMA model is to capture the complete range of human emotions, and the dynamics of emotions, for instance coping. The EMA model is the first computational models to expand its view from only emotions to another affective process: coping. EMA, unlike ALMA, does not use the OCC model or the PAD theory as a foundation to their implementation. They build their model on the theory of Smith et al. (1990). This theory falls in the category appraisal theories. One important feature of this theory that EMA also uses is the distinction between deliberative and automatic appraisal processes. With the deliberative appraisal process, the human is conscious of making the appraisal, while with the automatic process, it is not. Another aspect of the theory of Smith et al. (1990), is the distinction between two processes: appraisal and coping. Coping is usually related to stress, but Lazarus (1991) applied it to emotions too. Coping is a process to decrease (mainly) negative emotions. EMA models these concepts of appraisal and coping in five steps: (1) construct a interpretation of ongoing events in terms of beliefs, desires, and intentions, (2) generate appraisal frames with appraisal variables, (3) map the frames to instances of emotions, (4) incorporate the emotions into the current state, and (5) adopt a coping strategy. EMA, like ALMA, also includes mood in its model. The definition of mood they use is the same as Lazarus (1991) uses, moods are spread over a longer period of time, and are not related to a single object or event. Mood is computed by combining all the emotions belonging to the current state the intensities of all each emotion types (e.g. Joy, Fear). All the values are then passed through a sigmoid function, which results in a value between zero and one. This value will influence which emotion will have the largest intensity by adding intensity to the emotion close to the value of the mood. The appraisal process of emotion in EMA is more detailed than in ALMA: it is divided in more separate processes, and more variables are included. EMA also includes a coping process, ALMA does not have a way to reevaluate its emotions. Mood is also included in EMA, but plays a small role. It merely decided which emotion is strongest at a time, while in ALMA the mood has influence on all the emotions, and on future emotions.

2.3.2 Logical structure of emotion

The first one to introduce a purely logical model of emotion was Meyer (2004). He aimed to describe how the emotional states of emotional agents evolve over time. This is accomplished by extending the KARO framework (a blend of dynamic and epistemic/doxastic logic) with four basic emotions: happiness, sadness, anger, and fear. Steunebrink et al. (2007, 2009); Steunebrink (2010) built a more extensive logic on this idea, using the OCC model as inspiration for structuring multiple

emotions. In Steunebrink (2010) a complete formalization of the OCC model is proposed. This is done to build a bridge between the psychological literature and the implementations of emotions in robots and virtual characters, the formalization provides a foundation for computational models. The first important step to formalization in Steunebrink (2010) is providing a logical structure for the emotion elicitation process. As shown above in figure 5, Ortony et al. (1988) already provided a structure to their emotions. However, this structure is not compositional yet, and since logic is compositional, the structure of the OCC model should be too. Therefore Steunebrink (2010) proposes the structure seen in figure 8. This structure can be seen as an inheritance structure: the emotion types are specifications of those above them, and generalizations of those below them. Like the OCC model, at the top, the distinction can be made between positive and negative emotions. Next they can be specified by whether they are events, actions or objects, and respectively by pleased/displeased, approving/disapproving, and liking/disliking. Aspects of objects leads to roughly the same emotions as the OCC model, except that it does not specify further to love and hate. Actions of agents have the same structure as the OCC model. The main difference occurs with consequences of events: whether an event is prospective here only can lead to hope and fear, instead of to hope, fear, satisfaction, fears-confirmed, relief, and disappointment in the OCC model. In Steunebrink (2010) either joy or distress leads to those emotions.

The formalization of this structure is done by first formalizing the structure in logic, then putting most concepts in dynamic doxastic logic, and finally grounding the work in KARO (Meyer, 2006; Meyer et al., 1999). An important note here is that this is the first part of the formalization; the 'triggering' of emotions. How they are experienced and regulated is a different process. An overview of the complete process of emotion, as seen by Steunebrink (2010), is given in figure 9. Triggering of the emotions is part of the appraisal process, and does not guarantee that emotions will also be experienced.

The logical structure is formalized by starting with on top of the structure, and specializing the emotions more while going lower in the structure. The eliciting conditions of the first node in the structure; positive and negative, are formalized as follows:

$$Positive_i^T(X) = Perceive_i \wedge Good_i(X)$$
$$Negative_i^T(X) = Perceive_i \wedge Bad_i(X)$$

$$Perceive_i(X) = PerceiveConseq_i(X) \vee PerceiveAction_i(X) \vee PerceiveObject_i(X)$$

The X stands for the situation the emotion is respective to. The superscript T indicates that we are talking about triggered emotions, not experienced emotions. The subscript i means agent i is involved. We can read $Positive_i^T(X)$ as 'agent i has a triggered positive feeling about X'. $Perceive_i(X)$ can be read as 'agent i perceives X'. Agents can perceive three types: events, actions, and objects. Good

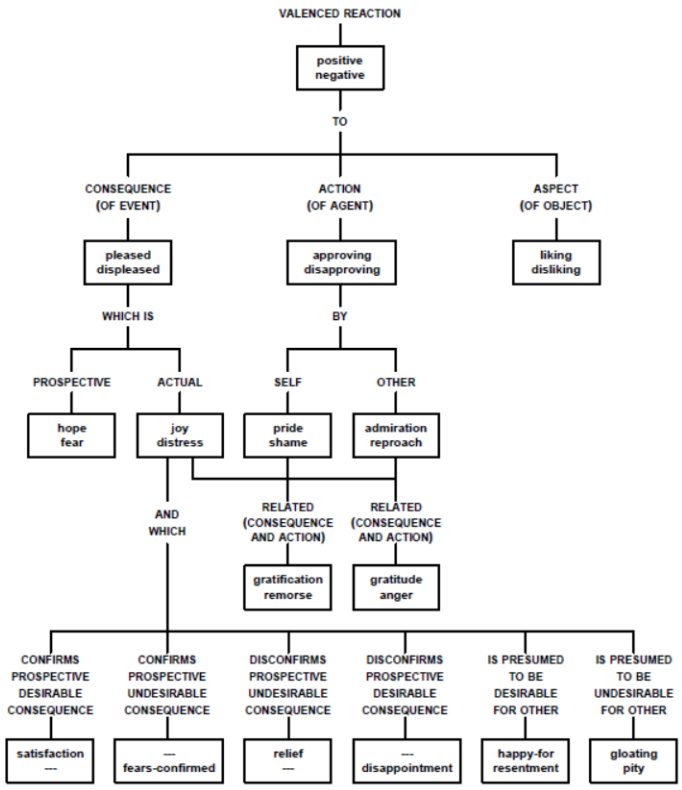

Figure 8: An inheritance based view of the eliciting conditions of emotions of the OCC model. Taken from (Steunebrink, 2010, p.31)

and Bad is divided in six operators:

$$Good_i(X) = Des_i(X) \vee Praisew_i(X) \vee Appeal_i(X)$$
$$Bad_i(X) = Undes_i(X) \vee Blamew_i(X) \vee Unappeal_i(X)$$

Des stands for desirable, Praisew for praiseworthy, Appeal for appealing, Undes for undesirable, Blamew for blameworthy, and Unappeal for unappealing. Now the distinction between events, actions, and objects are formalized in the following way:

$$Pleased_i^T(c) = PerceiveConseq_i(c) \wedge Des_i(c)$$
$$Displease_i^T(c) = PerceiveConseq_i(c) \wedge Undes_i(c)$$

$$Approving_i^T(j:a) = PerceiveAction_i(j:a) \wedge Praisew_i(j:a)$$

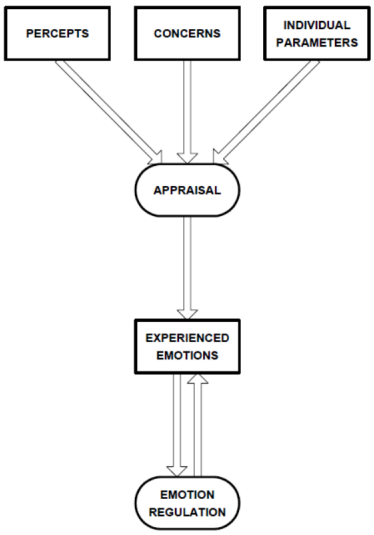

Figure 9: The procedural view of emotions: square boxes are data, round boxes processes. Taken from (Steunebrink, 2010, p.8)

$$Disapproving_i^T(j:a) = PerceiveAction_i(j:a) \wedge Blamew_i(j:a)$$

$$Liking_i^T(x) = PerceiveObject_i(x) \wedge Appeal_i(x)$$
$$Disliking_i^T(x) = PerceiveObject_i(x) \wedge Unappeal_i(x)$$

These definitions formalize the general emotion types at the top of the structure, seen in figure 10.

Joy and Distress are formalized as follows:

$$Joy_i^T(c) = Pleased_i^T(c) \wedge Actual_i(c)$$
$$Distress_i^T(c) = Displeased_i^T(c) \wedge Actual_i(c)$$

Where $Actual_i(c)$ can be read as 'agent i considers c to be an actual consequence of an event' (Steunebrink, 2010). The other emotions are formalized in a similar way, and can be found in appendix B.

The emotion eliciting conditions are formalized as shown above, the experience of emotions, its intensity and other factors are not included in this. A more detailed version of the structural view of emotions as was shown in figure 9 is given in figure 11. We can see that the emotion elicitation structure is part of appraisal, and experience comes after this appraisal. The experience of emotions contains one process and two data blocks. Whether an emotion is experienced depends on the

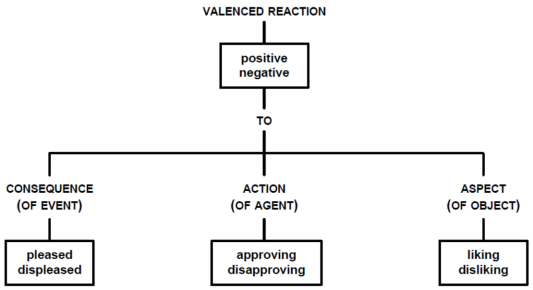

Figure 10: The upper part of the structure shown in figure 8

intensity of the emotion: if the intensity is high enough it will be experienced. This intensity is calculated with its potential, a threshold, and mood. The potential is the intensity the agent gets from the triggered emotions, the threshold determines when the emotion is experienced and is dependent on mood. Mood is a real number, and is subtracted from the threshold when the emotion is positive and added to the threshold when the emotion is negative.

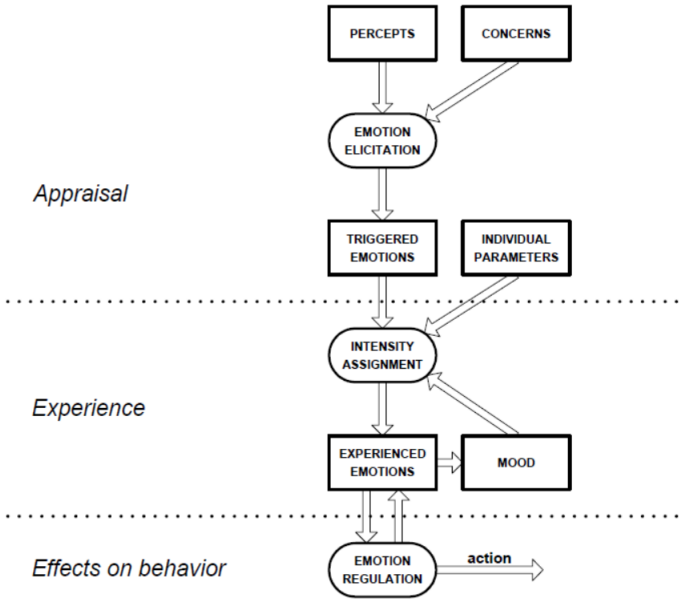

Figure 11: More detailed version of the procedural view of emotion. Taken from (Steunebrink, 2010, p.106)

The logical model of Steunebrink (2010) gives a complete account for the elici-

tation of emotions, that is, the appraisal part. The experience part contains some uncertainties and unexplained parts, as will be shown in the next chapter.

Around the same time as the logical model of Steunebrink (2010), a similar model was developed by Adam et al. (2009). This model also aims to provide a faithful formalization of the OCC model, and also does this with BDI-based logic. However, they made some assumptions that makes the logic quite different, especially in what it can do. For instance, Adam et al. (2009) assume that desires are free of contradictions, 'that agents have complete introspection', and that all actions are deterministic. Steunebrink (2010) does not make these assumptions, therefore it might explain more psychological phenomena, e.g. having mixed feelings about something. In the emotion elicitation process, the concept of a triggered emotion that is not experienced yet is important. Adam et al. (2009) do not make the distinction between triggered and experienced emotions. For instance, Joy is defined as $Joy_i\phi = Bel_i\phi \wedge Des_i\phi$. With this definition it is unclear whether joy is triggered or experienced.

2.4 Conclusion and first answers to research questions

In this chapter we saw that there are many different theories in different fields on the subject of emotion and mood. There are differences, but also common factors. With this information we can answer the first two subquestions. They are repeated below.

1. What are the affective phenomena *mood* and *emotion* exactly, how do they differ, and what is their relation?

2. How are *emotion* and *mood* formalized in agent models?

In sections 2.1.1 and 2.1.2, it became clear that emotion is a reaction to something in the environment of the subject. What this reaction is exactly, differed per theory. In this thesis, we will stay closely to the definition from the OCC model: emotions are valenced reactions to events, agents, or objects. This definition makes a clear enough distinction that it is possible to model emotions with it. The definition is not of great importance though, it provides some clarity but otherwise the exact definition of emotion is also defined by the theory or model. Mood is different from emotion in both duration and the source of the mood. It is agreed widely that moods have a longer duration than emotions, in both psychology and AI. It is also agreed on that mood does not originate from a situation, it emerges from someones mind. In this thesis we agree with the idea of Mehrabian (1980) that mood is an average of emotions over a certain time period. The relation between mood and emotion is a close one, they influence each other. Mood can influence the intensity of the emotions, and emotions influence what mood will occur.

Emotion is formalized in many computational models, in this thesis we investigated two of them that also have additions like mood (ALMA) and coping (EMA). Emotions are formalized in different ways, this has to do with the theory that the model is built on. ALMA uses the OCC model for emotion, EMA uses the theory

from Lazarus (1991). In the logical models we have looked at, emotions are formalized according to the OCC model. An important aspect of the logical model of Steunebrink (2010), was the modularity of the structure: the structure had to be build in such a way that it could be seen as an inheritances structure, this way it could be adopted in logic. Mood is formalized in ALMA, and a suggestion for a formalization of mood is given in Steunebrink (2010). In ALMA, the PAD theory was used, dividing mood in eight different kinds of mood according to their octant. The emotions from the OCC model were given PAD values such that a mood could be calculated from them. In Steunebrink (2010), mood could be either positive or negative, and influence the threshold of an emotion for being experienced. The exact calculations and logic were not provided.

Now that we have a clearer understanding of *emotion* and *mood*, and how it can be used, we will try to integrate mood with the logical model of emotion from Steunebrink (2010) next chapter. In the next chapter we will answer the last subquestion.

3 A conceptual integrated model of mood and emotion

In the previous chapter we have given the necessary background to understand emotion, mood, their interaction and their relation to artificial intelligence. We have also answered the first research questions, which are the building blocks to our main research question. Now we can use this information to solve the other research question, and with those the main question. The questions are repeated below.

- How can mood be integrated in models of agents with emotions?

1. What are the affective phenomena *mood* and *emotion* exactly, how do they differ, and what is their relation?
2. How are *emotion* and *mood* formalized in agent models?
3. How can *emotion* and *mood* be integrated?

We saw that mood and emotion are ambiguous concepts, and concluded that the definition of emotion and mood that we are going to use in this thesis will be one that is more directed towards the function of the affects. This is also the approach that most agent models with emotions took. If we would want to define an exact definition of emotion, the definition from the OCC model best suits our purpose: Emotions are valenced reactions to events, agents, or objects. An exact definition of mood is even harder to find, so we want to describe moods by their main properties: Moods last longer than emotions, and are not specified towards an event, action, or object. We saw several representations of mood and emotion in formal models, and in most, more attention is spent on emotion than on mood.

In this chapter, we are going try to find a way to integrate mood with the work of Steunebrink (2010). We believe this work lends itself best for integrating mood, since it has a minimal amount of assumptions and is written in a way that adding parts is not disrupting for the model. In section 3.1, we will specify the need for such a change to the model. We will go more into detail about how mood is used in Steunebrink (2010), and discuss what could be improved in this model. Then in section 3.2, we will discuss how we can integrate mood with the model, and what is needed for it to work. We will then propose a new structure to accommodate an integration with mood. Lastly, section 3.3 will treat the first necessary transformation that the model must undergo to support mood.

3.1 Motivation: a closer look at mood in an agent model with emotions

In the previous chapter we saw a short overview of the logical structure of emotions by Steunebrink (2010). This work is mostly on emotion, but the importance of mood in the emotion process is not neglected. In fact, moods are used to influence

the threshold of emotion experience in a negative or positive way. In other words, mood influences how easily an emotion is experienced. To see how this is done, we repeat the process of emotion as used in this work below in figure 12.

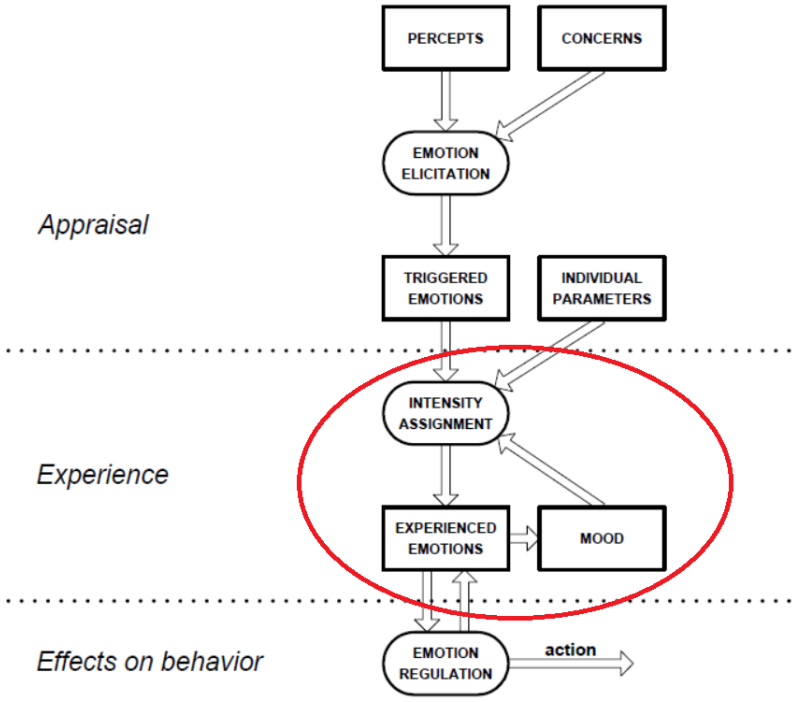

Figure 12: The procedural view of emotion according to Steunebrink (2010), repeated figure

The red circled part of the figure is where mood influences emotion. Rounded boxes stand for processes, square boxes for data. In the appraisal part, percepts and concerns serve as data that is used in the emotion elicitation process. The emotion elicitation process is a detailed process that distinguishes between events, actions and objects. We saw a structural view of this process in figure 8. The outcome of this process are triggered emotions. Together with 'individual parameters' and mood, they are the input for the process of the intensity assignment in the experience part of the structure. The output of the intensity assignment are experienced emotions. These experienced emotions are input for emotion regulation and for mood. Mood is here shown as a data block, so there are no processes involved in determining the mood. Steunebrink (2010) suggest that mood is calculated by averaging over intensities of recently experienced emotions. This will result in an overall mood, which will influence the threshold of the initial intensity

of an emotion. They also state that over time, mood should return to zero so that it is self stabilizing.

There are several problems with this approach of formalizing mood. First, there is no explanation as to how mood is calculated, it is only speculated that 'mood is calculated based on the intensities of recently experienced emotions' (Steunebrink, 2010, p.114). We can see this problem in figure 12, mood is depicted as just data, while there should be processes involved to determine how the mood is calculated. The second problem is that there is no theory of how the mood behaves over time, it is stated that mood should be self-stabilizing, that is, it should return to value zero over time when no new emotions are triggered. How this should be accomplished is unclear. Lastly, in this logical model, mood can be either positive or negative, which, one could argue, is a too simplistic view of mood. Steunebrink (2010) admit themselves that mood is multidimensional in humans, but they assume that a unidimensional mood can always be determined. As we saw in ALMA, a multidimensional mood can be realized in a formal way by using PAD values.

In some other formal models that we mentioned in the previous chapter, mood is completely left out (e.g. Adam (2007); Gratch and Marsella (2004). In ALMA however, mood plays a major role in the process of emotions. As we saw in section 2.3, mood influences the intensity of emotions. For instance, when an agent is in an exuberant mood (which is Pleasurable, Aroused, and Dominant), negative emotions like distress will have a lower intensity then when the mood was neutral. This is because the mood 'pulls' the PAD values of the emotion towards it, so strongly negative PAD values will become more mildly negative because of the mood having positive PAD values. Mood was here depicted as a medium-term affect with three dimensions (pleasure (P), arousal (A), and dominance (D)) (Gebhard, 2005). Since ALMA has a detailed account of how mood is incorporated in their emotion process, and the mood is depicted as multidimensional instead of unidimensional, we will explore the options of combining the idea of mood as used in ALMA with the logical model of Steunebrink (2010) in the next section.

3.2 The process of emotion and mood: a new structure

To integrate mood and emotion in the logic of Steunebrink (2010), we need a concept of mood which does justice to the complexity of mood, but which is also precise. In ALMA (Gebhard, 2005), such a definition is used. They described mood in such a way and with such functions that it lends itself excellently for formalization, as we saw in the previous chapter. We can find inspiration in ALMA to improve the concept of mood in a logical way and to integrate it with emotion.

In ALMA, the process of an emotion coming into existence is different from the process that Steunebrink (2010) proposed. We saw in the previous chapter that the process from ALMA looks like the one in the figure 13.

Processes are shown in ellipses and data in square boxes. We can roughly divide this process in three parts: Data organizing (top level), data manipulation

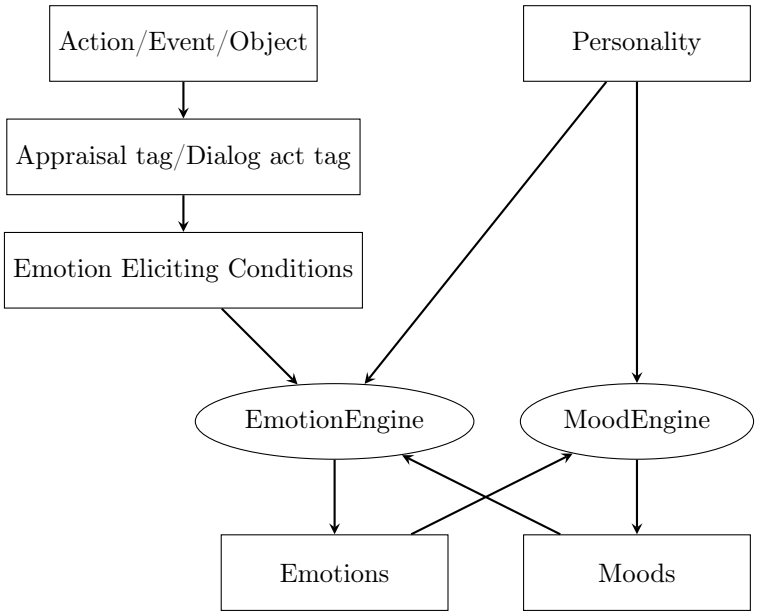

Figure 13: An overview of the process of emotion and mood in ALMA, repeated figure

(Emotion Engine and Mood Engine), and results (bottom level). The top level with the data must be defined in XML like files before the computations starts and serve as input for the Engines. The steps from action/event/object to EEC are called appraisal in ALMA. Next moods and emotions are computed with the *MoodEngine* and the *EmotionEngine*. The results, emotions and moods, are then again used as input for the engines. (Gebhard et al., 2003a, 2004; Gebhard, 2001; Gebhard et al., 2003b). The MoodEngine and the EmotionEngine are the central parts in this structure where the emotions and moods are calculated. The essential psychological theories that underlay this implementation of mood are the use of the PAD values (Mehrabian, 1996b), and the Big Five model of personality (McCrae and John, 1992). Mood is placed in a three dimensional space with these PAD values, as explained in chapter 2.

To incorporate this idea of mood from ALMA with the logical model of Steunebrink (2010), we need to make several adjustments to the process of emotion and the structures in the logical model. We want emotions to be able to have PAD values, and we also want a way to incorporate personality. In ALMA, mood is defined with PAD values, and calculated through emotions. The emotions in ALMA also have PAD values to be able to use them for the calculation of the mood. So to be able to use mood in the logical model of Steunebrink (2010) as it is used in ALMA, we need to add PAD values to the emotions that Steunebrink (2010) uses. The emotions need to have PAD values because the mood is calculated from the

emotions, and mood is defined with PAD values. The MoodEngine in ALMA not only takes emotions as input, but also personality. This should also be added to the logical model. These additions are hard to incorporate in the original structural view of emotion that was shown in figure 12. Therefore, before we explore how to execute these adjustments, we are going examine the differences between the structures of the emotion process in ALMA and in Steunebrink (2010), and we are going to look into an alternative structure for the process of emotion (and mood).

If we look at the general structure of the process of emotion as depicted by Steunebrink (2010) in figure 12 next to the structure of ALMA (Gebhard, 2005) in figure 13, we can see some similarities and differences. In ALMA, the steps from the event to the EEC are called appraisal. This appraisal differs a lot from the appraisal part in the logical model: In the first (ALMA) events are simply given 'tags' to determine the emotion, in the second (the logical model), the appraisal part consists of the complex inheritance structure which defines conditions in which the emotion is triggered (see section 2.3.2). The 'experience' part from the logical model is not comparable with the components of the ALMA model since there is no distinction between the emotion being elicitated and being experienced in ALMA. In ALMA, Mood is a process (the MoodEngine) that interacts with the emotions. In the structure of Steunebrink (2010), Mood is just data, and it interacts with intensity and experienced emotions. The structure of the emotion process in the logical model makes it hard to add PAD values, to add personality, and to let mood interact with emotions as they do in ALMA. This is because mood has no influence on the triggered emotions in its current placement, it only influences the intensity. In ALMA, mood influences all the emotions, both triggered and experienced. We propose a new structure in figure 14 where mood interacts with emotions, and where we use percepts and concerns as input for the emotion process. We think that this way we can make minimal changes to the logical model and still allow for an easy integration with mood.

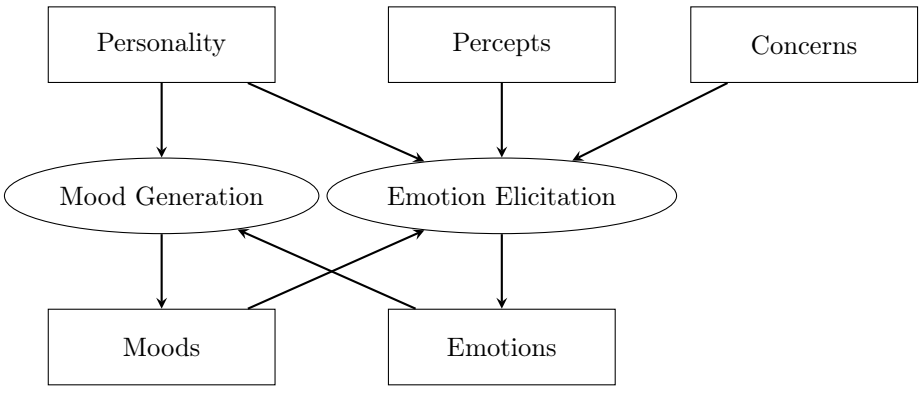

Figure 14: A structure with mood added in the appraisal part of the procedural view

In the new structure in figure 14, we have three data blocks that are input for mood and emotions processes. We have percepts and concerns, used in the same way as in Steunebrink (2010). Personality contains the personality of the agent according to the five-factor model of McCrae and John (1992). Then below these data blocks, we have two processes: Mood Generation and Emotion Elicitation. The first calculates moods based on current emotions and the personality. Emotion Elicitation calculates emotions using percepts, concerns, personality, and the current mood the agent is in. This structure allows for the interaction between mood and emotions, and adds the personality of an agent. When we compare this new structure to the structure in figure 12, we can see that we no longer make the difference between triggered and experienced emotions. There also appears to be no intensity assignment anymore. We chose to leave the intensity assignment process out, since it is not complete in the theory of Steunebrink (2010), and it is also in an implausible place in the structure, namely, the experience part. In the OCC model, intensities come from the knowledge, goals, and beliefs of the agent Ortony et al. (1988). Therefore the intensities are already determined in the Percepts and Concerns data blocks in figure 14.

This new structure of the emotion process will allow us to adopt the ideas for mood from ALMA more easily. In the next section we are going to zoom in and discuss the first thing we would need to be able to add mood as used in ALMA, to the logical model of Steunebrink (2010): combining the emotions from the OCC model with the PAD values.

3.3 Adding PAD to Emotions

To add the definition of mood as used in ALMA to the logical model of Steunebrink (2010), we need to add PAD values to the emotions used in the logical model. ALMA uses the PAD values Pleasure, Arousal, and Dominance, to describe mood. They form a three dimensional space, where every octant of the space illustrates a different mood. To calculate these moods, ALMA uses emotions. They use the 22 emotion types from the OCC model, the same emotions as the logical model uses. To calculate the mood from the emotions, emotions need to have the same kind of values as moods: PAD values. The Mood Engine then can average over the PAD values of the current emotions, and it produces a mood. ALMA also needed to add PAD values to its emotions to make this calculation, since they used the emotions from the OCC model, the same emotions as the logical model of Steunebrink (2010) uses, which do not have PAD values. They proposed a mapping which maps the emotions from the OCC model to specific PAD values. This mapping is given below in table 3.

Since the logical model of Steunebrink (2010) uses the same emotions, we can use this mapping too to determine what the PAD values should be for the individual emotions. One might notice that there are two emotions in this mapping that are not used in Steunebrink (2010)s model, *hate* and *love*. These emotions can be seen as sub-types of *liking* and *disliking*.

We encounter several problems when we use this mapping to give emotions PAD

Emotion	**P**	**A**	**D**	**Mood Octant**
Admiration	0.5	0.3	-0.2	+P+A-D Dependent
Anger	-0.51	0.59	0.25	-P+A+D Hostile
Disliking	-0.4	0.2	0.1	-P+A+D Hostile
Disappointment	-0.3	0.1	-0.4	-P+A-D Anxious
Distress	-0.4	-0.2	-0.5	-P-A-D Bored
Fear	-0.64	0.60	-0.43	-P+A-D Anxious
FearsConfirmed	-0.5	-0.3	-0.7	-P-A-D Bored
Gloating	0.3	-0.3	-0.1	+P-A-D Docile
Gratification	0.6	0.5	0.4	+P+A+D Exuberant
Gratitude	0.4	0.2	-0.3	+P+A-D Dependent
HappyFor	0.4	0.2	0.2	+P+A+D Exuberant
Hate	-0.6	0.6	0.3	-P+A+D Hostile
Hope	0.2	0.2	-0.1	+P+A-D Dependent
Joy	0.4	0.2	0.1	+P+A+D Exuberant
Liking	0.40	0.16	-0.24	+P+A-D Dependent
Love	0.3	0.1	0.2	+P+A+D Exuberant
Pity	-0.4	-0.2	-0.5	-P-A-D Bored
Pride	0.4	0.3	0.3	+P+A+D Exuberant
Relief	0.2	-0.3	0.4	+P-A+D Relaxed
Remorse	-0.3	0.1	-0.6	-P+A-D Anxious
Reproach	-0.3	-0.1	0.4	-P-A+D Disdainful
Resentment	-0.2	-0.3	-0.2	-P-A-D Bored
Satisfaction	0.3	-0.2	0.4	+P-A+D Relaxed
Shame	-0.3	0.1	-0.6	-P+A-D Anxious

Table 3: Mapping of OCC emotions to PAD values, taken from (Gebhard, 2005, p.4)

values in the logical model. As mentioned before, absolute values are seldom used in logic, and therefore we cannot copy the exact PAD values of these emotions and calculate the mood from it in the same way as ALMA does. What we can do, is use this mapping to determine in which mood octant the emotion falls. Remember from the previous chapter that mood could be divided in eight different moods, each one matching with one of the octants in the three dimensional PAD space. If we can determine the octant in which an emotion falls, we do not have to know the exact point of the emotion, which is the exact PAD value. This way we can calculate mood based on the octants of the emotions instead of the exact points in the PAD space. The PAD values of the emotions without the exact values would then look like the PAD values of the moods. For instance, the PAD value of Joy would be +P, +A, and +D. The mood that belongs to joy then has the same PAD value: +P, +A, and +D. We can express the octants of the PAD space with the different moods and their emotions with six values: Pleasant (+P) and Unpleasant

(-P), Aroused (+A) and Unaroused (-A), and Dominant (+D) and Submissive (-D). Now that we have eliminated the exact PAD values of the emotions, we can use the mapping of ALMA in figure 3 to determine which emotions falls into which mood octant. In table 4 we organized the emotions according to their mood, based on the mapping in table 3. Each emotion has the same PAD value of the mood that its in.

Mood	**Emotions**			
Exuberant (+P+A+D)	Gratification	Happy-For	Joy	Pride
Dependent (+P+A-D)	Admiration	Gratitude	Hope	Liking
Relaxed (+P-A+D)	Satisfaction	Relief		
Docile (+P-A-D)	Gloating			
Hostile (-P+A+D)	Anger	Disliking		
Anxious (-P+A-D)	Disappointment	Fear	Shame	Remorse
Disdainful (-P-A+D)	Reproach			
Bored (-P-A-D)	Distress	Fears-Confirmed	Pity	Resentment

Table 4: Emotions sorted by their moods according to PAD values given in ALMA (table 3)

The mood Exuberant (Pleasant (+P), Aroused (+A), and Dominant (+D)) has four emotions: Gratification, Happyfor, Joy, and Pride. Dependent (Pleasant (+P), Aroused (+A), Submissive (-D)) has also four emotions: Admiration, Gratitude, Hope, and Liking. Relaxed (Pleasant (+P), Unaroused (-A), and Dominant (+D)) has two emotions: Satisfaction and Relief. Docile (Pleasant (+P), Unaroused (-A), and Submissive (-D)) has only one emotion: Gloating. Hostile (Unpleasant (-P), Aroused (+A), Dominant (+D)) had two emotions: Anger and Disliking. Anxious (Unpleasant (-P), Aroused (+A), Submissive (-D)) has four emotions: Disappointment, Fear, Shame, and Remorse. Disdainful (Unpleasant (-P), Unaroused (-A), Dominant (+D)) has one emotion: Reproach. And lastly Bored (Unpleasant (-P), Unaroused (-A), Submissive (-D)) has four emotions: Distress, Fears-confirmed, Pity, and Resentment.

Now we have given PAD values to emotions on a surface level, but integrating the PAD values with the emotions in the logical model is still not an easy task, even if we only work with positive and negative PAD values. We have to keep in mind that the emotions are elicited in a process, to be more precise, they are elicited in a compositional way through the structure Steunebrink (2010) provided (figure 8). This elicitation structure has multiple levels: on the top the positive/negative distinction, then a division in events, actions, and objects, and so on. Lower in the structure different emotions are born from different conditions, and sometimes from other emotions: for instance the emotion *gratification* is the combination of the emotions *joy* and *pride*. This elicitation structure is compositional and therefore the lower nodes inherit the features from the nodes above them. So where in this emotion elicitation process can we add PAD values to the emotions? The obvious

way would be attaching a PAD value to each separate emotion. This might not be possible however, since some emotions are below other emotions in the structure. This would mean that the lower emotion inherits the features, so also the PAD values, from the higher emotion. For instance, *gratitude* is combination of two other emotions: *joy* and *admiration*. *Gratitude* would therefore inherit the PAD values of joy and admiration, which might not lead to the right PAD values for gratitude. To see how we can add the PAD values to the emotions, we are going to analyze the emotion elicitation structure of Steunebrink (2010), the moods that belong to the emotions, and the PAD values of the emotions, and try to find a pattern.

In figures 15 and 16 at the end of this chapter, we see the emotion elicitation structure of the logical model of Steunebrink (2010). The structures are already split by the upper node: the positively or negatively valenced reactions. For the complete structure, see figure 8. We gave each emotion the color of the mood the emotion is in according to the PAD values of the mapping from ALMA: red for Exuberant, pink for Dependent, orange for Relaxed, yellow for Docile, blue for Anxious, green for Hostile, light blue for Bored, and purple for Disdainful. This way we can see the patterns between related emotions and their moods more easily.

One thing that springs to the eye when looking at these figures, is the similarity between the positive/negative distinction and the P value. All the emotions and moods with a positive P value, that is +P, are positively valenced reactions. The emotions with positive P values are exactly the same emotions that are valenced as positive in the logical structure, and vice versa for emotions with a negative P value. This means that we can add the P value to the emotions at the top of the process in the same place as the positive/negative distinction, since the same emotions are differentiated by the P value in the PAD theory and by the positive/negative distinction.

The other values, Aroused (+A), Unaroused (-A), Dominant (+D), and Submissive (-D), are harder to pin down. We can see some patterns if we look closely at the positively valenced emotions. For instance, we can see a pattern in figure 15: actions by an agent itself lead to the emotions pride and gratification, and both have the mood Exuberant (red). For the A and D values, this means that positive actions by an agent itself have a +A and a +D value. Positive actions by another agent lead to the emotions admiration and gratitude, and both have the mood Dependent. For the A and D values this means that actions by another agent have a +A value and a -D value. Because the lower emotions like gratification and gratitude have the same mood as those above them, the compositionality of the structure stays intact. In this case we could add a +A and +D value to the emotion pride, and +A and -D to admiration, and the emotions below them would inherit those values and be in the right mood at the same time. However, when we look at the negatively valenced reactions (-P) in figure 16, taking this approach would violate the compositionality. Negatively valenced actions by an agent itself lead to the emotions shame and remorse, which both have the mood Anxious, and have a +A value and a -D value. Here we could add the A and D values to shame, and remorse would inherit the values without complications. But if we look at

actions by another agent, this approach is no longer viable: reproach has the mood disdainful (-A, +D), and anger has the mood hostile (+A, -D). Because of the compositionality, we would expect anger to inherit the mood and the A and D values from reproach. This is not the case, both values change. We can also easily see this because the color of the mood changes from reproach to anger. Adding the A and D value to reproach would therefore not be an option since anger cannot inherit its values. When looking at other emotions in the structure, for instance those under *joy* and *distress*, it is even harder to see a pattern. After closer inspection, the only pattern that seems consistent is the prospective consequence of an event: *hope* and *fear*. These emotions have a Dependent and Anxious mood respectively, and always have Aroused (+A) and Submissive (-D) as A and D values. Because of the discrepancies in the structure, we cannot add the A and D values simply to an emotion or an to eliciting condition.

This emotion elicitation structure is helpful for seeing patterns regarding the compositionality of the structure, because we can see directly if an emotion 'inherits' the same mood and PAD values as its parent by checking if the color of the mood changes lower in the structure. However, there might be other patterns that are not as visible. In this structure we only see +P and -P separately (15 and 16 respectively) and the A and the D value are always together. In the structure we immediately saw a useful relation: the relation between the positively or negatively valenced reaction and the P value. This relation is only visible because the structures are separated by their P value. Therefore we also want look at A and D values of the emotions and their moods separately, to see if there are more patterns like this. By looking at the A and D values separately, we can also see more clearly what their relation is to the P value. Table 5 shows this analysis of the separate A and D values.

The table is read as follows: the first row indicates whether we are looking at the emotion with Pleasant (+P) or Unpleasant (-P), twice because of the A and D value. The +P value indicates the A and D value of the positively valenced emotion, the -P of the negatively valenced emotion. We look at the emotions compared to their +P or -P value because the logical structure also makes the distinction with positively and negatively valenced reactions. This way we can also see what the relation of the A or D value is to the P value, and therefore we can see relations between eliciting conditions and the A and D values more clearly. The first column shows the emotions and their eliciting conditions. After the emotions their A and D value is given according to the P value. For example, the second row is read as: 'The emotions hope and fear, which are prospective events, have an Aroused (+A) value for both a +P (hope) and a -P (fear) value, and an Submissive (-D) value for both +P (hope) and -P values (fear)'. It can also be read more shortly as 'hope has -D and +A, fear has -D and +A'. The first division in the eliciting structure of the emotions, the division between events, actions, and objects, is indicated with a double horizontal line. So from top to bottom we first see the emotions that originate from events, then those that come from actions, and lastly the emotions

Emotions and eliciting condition	**+P**	**-P**	**+P**	**-P**
hope/fear: event - prospective	+A	+A	-D	-D
joy/distress: event - actual	+A	-A	+D	-D
satisfaction: event - actual + confirm desirable	-A	–	+D	–
fears-confirmed: event - actual + confirm undesirable	–	-A	–	-D
relief: event - actual + disconfirm undesirable	-A	–	+D	–
disappointment: event - actual + disconfirm desirable	–	+A	–	-D
happy-for/resentment: event - actual + presume desirable	+A	-A	+D	-D
gloating/pity: event - actual + presume undesirable	-A	-A	-D	-D
pride/shame: action - self	+A	+A	+D	-D
admiration/reproach: action - other	+A	-A	-D	+D
gratification/remorse: action - self + consequence	+A	+A	+D	-D
gratitude/anger: action - other + consequence	+A	+A	-D	+D
liking/disliking: object	+A	+A	-D	+D

Table 5: Emotions and their eliciting conditions and their separate PAD values. Green cells indicates independence of P, red cells indicates a inverse relation with P

that originate from objects. Special cases in the table are the emotions that are in the lowest part of the elicitation structure, these emotions do not always have a positive and negative valenced counterpart. In these cases a line is placed where the counterpart value would be. Cells where the A or D value has an inverse relation with P are colored red. The inverse relation means that if the P value is positive, the A or D value is negative, and vice versa. Lastly we colored cells green that have an A or D value that is independent from the P value. This shows that the specific eliciting condition has influence on the A or D value, instead of a singular emotion. We will discuss these cases more carefully below.

In table 5 we can see the pattern we discussed earlier: *hope* and *fear* always have Aroused (+A) and Submissive (-D) values. Therefore we can say more about the PAD values of the eliciting conditions of these emotions instead of the PAD values of one specific emotion. In the case of hope and fear, we can say that events which are prospective (the eliciting conditions of these emotions) always have a +A and a -D value. We can also see patterns that we did not see before when looking at the structures in figure 15 and 16 with the colored moods. For instance when we look at gratification and remorse, which are elicited by the conditions 'action by the agent self and the consequence of an event', they always have a +A value, but their D value depends on the P value. In this case we can say that the eliciting conditions 'action by an agent self and the consequence of an event' leads to a +A value. The D value still depends here on the specific emotion. We can see the same pattern for the eliciting conditions of pride and shame, gratitude and anger, and liking and disliking. The eliciting condition of gloating and pity (events which are actual and presume undesirable for another agent), is the only other eliciting

condition beside hope and fear that leads to a set A and D value (-A and -D). However, this does not mean that we now add the A and D values to this emotion eliciting condition: gloating/pity is an elaboration on the eliciting conditions of joy/distress, but have other A and D values. When we would add the A and D value to the eliciting conditions, gloating/pity would have conflicting values since it also inherits the values from joy/distress. Therefore we cannot say that certain emotion elicitation conditions always lead to set PAD values.

In the white cells, the A and D values are the same as the P value above it. This means that the A and D values cannot be added to an emotion eliciting condition, but depend on the specific emotion. In the red cells, the A and D values are the opposite of the P value. In this case we also cannot add the A or D value to an emotion eliciting condition, since they are still dependent on the P value. We can see that less than half of the emotions and their eliciting conditions have A and D values that are independent of the P value. We will discuss this finding in the next chapter.

3.4 Conclusion

In this chapter we tried to add the concept of mood as used in ALMA to the formalization of Steunebrink (2010). Early on we ran into problems about the general process of emotion: The general structure of the process of emotion in the logical model was hard to unify with the structure used in ALMA. Therefore we introduced a new structure which is a hybrid between the two models. This new structure allowed us to add mood as it is used in ALMA to the logical model of Steunebrink (2010). The important part of an emotion coming into existence is the emotion elicitation process, which is the process that is worked out most carefully in the work of Steunebrink (2010). Mood in ALMA is defined by PAD values, therefore we concluded that the emotions in the logical model also should have PAD values in the elicitation process. ALMA provided a mapping for emotions from the OCC model that we used to categorize the emotions into moods. To see where exactly in the emotion elicitation process we should add the PAD values, we tried to find patterns between the emotions, the emotion eliciting conditions, and the PAD values. Since the structure of the emotion elicitation process is compositional, ideally we would want emotions lower in the structure to inherit its PAD values from its parents. Therefore emotions would have to be in the same mood as the emotion above it. In figures 15 and 16, we made the relations between PAD values, moods, and emotions, more insightful by giving a color to each emotion and its mood. We concluded that there were almost no clear patterns visible that would respect the compositionality. Therefore, we analyzed the A and the D values separately based on the P values (since there is a clear relation between the P value and the positive/negative distinction), and found that most values are dependent on P, and only a few where independent. This made us wonder how the PAD values were calculated in ALMA, and if we can find PAD values from the original PAD theory. Chapter 4 will treat these questions, and will also look critically at the compositional structure of the emotion elicitation process.

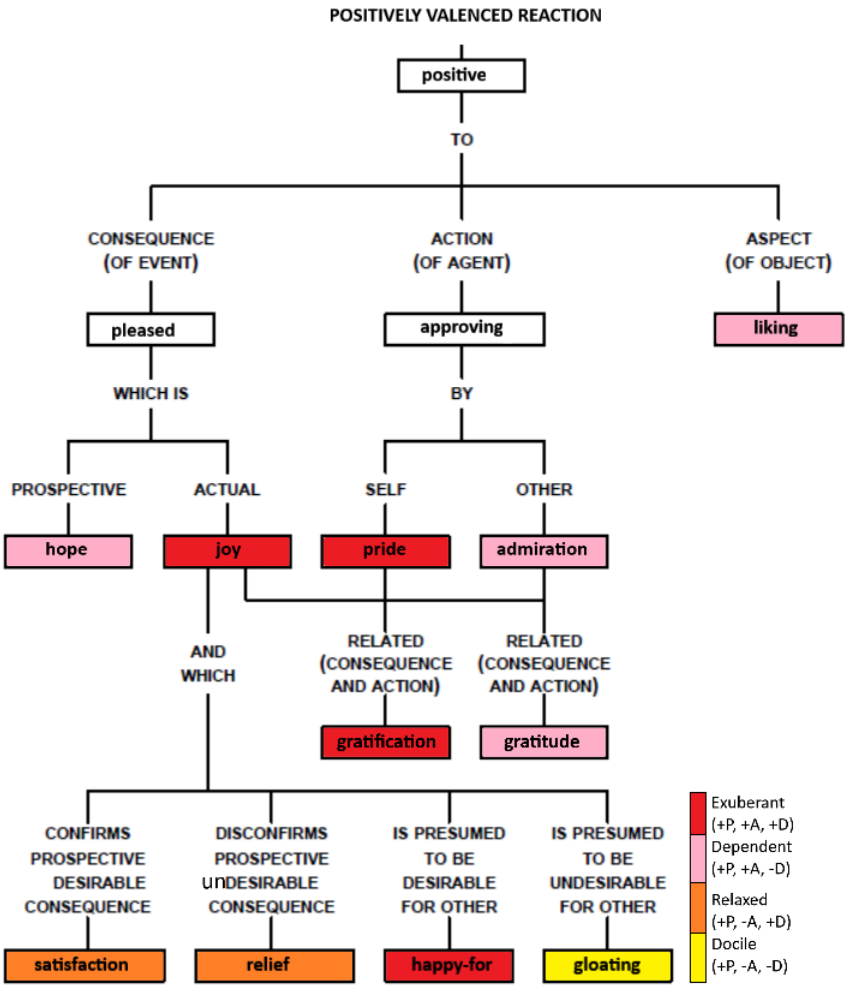

Figure 15: The emotion elicitation structure with moods added in colors to its positive emotions

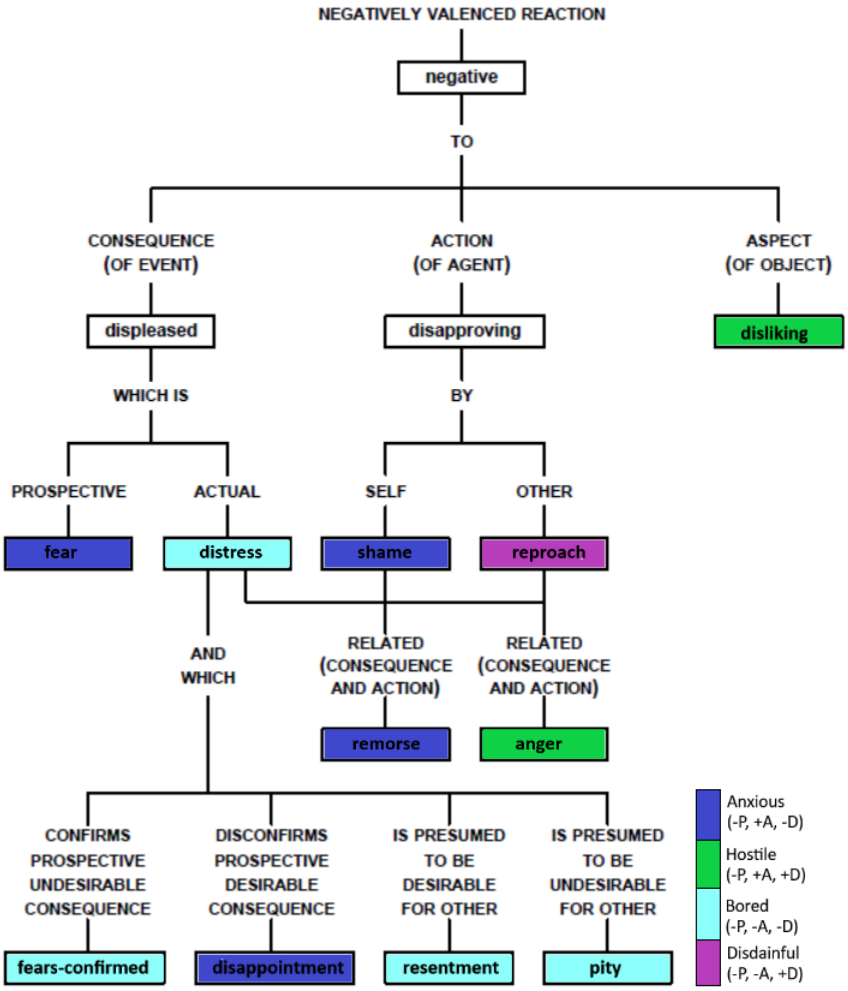

Figure 16: The emotion elicitation structure with moods added in colors to its negative emotions

4 Analysis

In the previous chapter we tried to integrate mood with the logical model of Steunebrink (2010) by changing the structure of the process of emotion, and by adding PAD values to the emotions in the emotion elicitation structure. We concluded that it would be hard to add PAD values to the emotions because the compositionality of the emotion elicitation structure is violated. Until now we have used the PAD values from ALMA, and the emotion elicitation structure from Steunebrink (2010). In this chapter we are going to look at the original PAD values (Russell and Mehrabian, 1977) and at the original emotion elicitation structure (Ortony et al., 1988), to see if we can make the integration of mood smoother.

4.1 The PAD values of OCC emotions

In the previous chapter we saw that the ALMA model provided a mapping for the emotions of the OCC model, giving them PAD values. When we used these simplified values, we saw no obvious pattern occur in the inheritance structure of the logical model of Steunebrink (2010). However, when separating A and D values and looking at their relation with the P value, we saw some regularities. These were shown in table 4.

When looking closely at the mapping provided in ALMA, there are a few things that stand out. First of all, there are four emotions that share their values with another emotion: *shame* and *remorse*, and *distress* and *pity*. According to the PAD theory, each emotion should have an original PAD value since the emotion is cognitively defined completely by these three dimensions (Mehrabian, 1980). These emotions therefore should not have the same PAD values. Second, the balance between the eight moods is not very even, the moods *docile* and *disdainful* both contain one emotion while for example *exuberant* contains four emotions. Lastly, some emotions have two decimal places, others have only one. These oddities prompted us to have a closer look at these values, how where they calculated? There was no information in ALMA about this, so we looked at the original values found through psychological experiments by Russell and Mehrabian (1977).

The original PAD values and their descriptors were obtained through several experimental studies, two of them were treated in chapter 2. The emotions of the OCC model that were used both in ALMA and in the logical model of Steunebrink (2010) often do not have an exact counterpart in emotions described in Russell and Mehrabian (1977), therefore we looked at the type identification (the label of the emotion), and at some tokens (the words that describe this emotion) of the OCC emotions. The label and the tokens are specified in the original OCC theory by Ortony et al. (1988), all the specifications of the emotions in the OCC model can be found in Appendix A. We used these labels and tokens to match the OCC emotion with emotions from the PAD theory. We tried to make the best match possible by providing an emotion that best matches the OCC emotion, as wel as providing other emotions that also match the emotion label or tokens of the OCC emotion.

In table 6 the results of this matching are summarized. The first column contains the OCC emotion label and its tokens (if there were any). The columns following it contain the PAD values as given in ALMA, and the mood that the emotion belongs to. The column under 'PAD emotions' shows emotions that were studied in Russell and Mehrabian (1977). We took emotions that fitted the emotion from the OCC model and its tokens as closely as possible. The columns after it show their PAD values and the mood. The last column displays the mood as it is called in Mehrabian (1980). We colored rows orange when there was a clash between the mood the emotion is in according the ALMA mapping, and the mood the emotion is in with the original PAD values. For example, *hope* is, according to the PAD values from ALMA, in the mood *Dependent*. When looking at the original PAD values of emotions that are close to the emotion hope, *hopeful* and *excited*, these emotions both are in the mood *Exuberant*. All rows that are colored orange will therefore have an effect on the previous patterns found in chapter 3 when the other mood is used. There are four PAD values from ALMA in fat print; these indicate that the same PAD value is used in another emotion (*shame* and *remorse*, and *distress* and *pity*).

Some emotions from the OCC model match better with emotions from the PAD theory than others; for example *joy* and its tokens have nearly equivalent emotions, whereas for *fears confirmed* (which has no tokens) there is no closely related emotion available. In these harder cases, we tried to match the OCC emotion with emotions that have related feelings, but it is important to keep in mind that because there is no clear equivalent, it is harder to draw conclusions on these emotions. In chapter 2, we also saw that there were discrepancies with some emotions in their study compared to previous studies. These emotions were *friendly*, *affectionate*, *aggressive*, *anxious*, and *distress*. We should take extra care when these emotions occur in our matching.

In table 6, we can see differences in the mood an emotion belongs to with the OCC emotions *happy-for*, *love*, *hope*, *liking*, *relief*, *gloating*, *disappointment*, *reproach*, *fears-confirmed*, *resentment*, *distress*, and *pity* (the rows colored orange). In the next section we will see what this means for our previous analysis, and later in the chapter we will also discuss the more unclear cases, for instance *fears-confirmed* and *distress*.

OCC emotions	P	A	D	Mood	PAD emotions	P	A	D	Mood	Mood PAD
Joy	0.40	0.20	0.10	Exuberant	Joyful	0.76	0.48	0.35	Exuberant	Exuberant
happy, joyful,					Happy	0.81	0.51	0.46	Exuberant	Exuberant
elated					Elated	0.50	0.42	0.23	Exuberant	Exuberant
					Enjoyment	0.77	0.44	0.42	Exuberant	Exuberant
Pride	0.40	0.30	0.30	Exuberant	Proud	0.77	0.38	0.65	Exuberant	Exuberant
Gratification	0.60	0.50	0.40	Exuberant	Self-satisfied	0.86	0.20	0.62	Exuberant	Exuberant
self-satisfaction, smug					Arrogant	0.00	0.34	0.48	Exuberant	Exuberant
Happy-for	0.40	0.20	0.20	Exuberant	Happy	0.81	0.51	0.46	Exuberant	Exuberant
delighted-for,					Respectful	0.38	0.13	-0.08	Dependent	Dependent
pleased-for					Appreciative	0.55	0.07	-0.14	Dependent	Dependent
Love	0.30	0.10	0.20	Exuberant	In love	0.82	0.65	-0.05	Dependent	Dependent
adore, affection, like					Loved	0.87	0.54	-0.18	Dependent	Dependent
					Affectionate	0.64	0.35	0.24	Exuberant	Exuberant
Admiration	0.50	0.30	-0.20	Dependent	Impressed	0.41	0.30	-0.32	Dependent	Dependent
appreciation, awe,					Awed	0.18	0.40	-0.38	Dependent	Dependent
respect					Respectful	0.38	0.13	-0.08	Dependent	Dependent
Gratitude	0.40	0.20	-0.30	Dependent	Thankful	0.61	0.10	-0.13	Dependent	Dependent
thankful, appreciation					Grateful	0.64	0.16	-0.21	Dependent	Dependent
Hope	0.20	0.20	-0.10	Dependent	Hopeful	0.51	0.23	0.14	Exuberant	Exuberant
excitement, hopeful					Excited	0.62	0.75	0.38	Exuberant	Exuberant
Liking	0.40	0.16	-0.24	Dependent	Affectionate	0.64	0.35	0.24	Exuberant	Exuberant
adore, affection, like					Appreciative	0.55	0.07	-0.14	Dependent	Dependent
					Friendly	0.69	0.35	0.30	Exuberant	Exuberant
Satisfaction	0.30	-0.20	0.40	Relaxed	Secure	0.74	-0.13	0.03	Relaxed	Relaxed
					Untroubled	0.79	-0.01	0.33	Relaxed	Relaxed
Relief	0.20	-0.30	0.40	Relaxed	Consoled	0.29	-0.19	-0.28	Docile	Docile
					Relaxed	0.68	-0.46	0.06	Relaxed	Relaxed
					Quiet	0.19	-0.40	-0.04	Docile	Docile
Gloating	0.30	-0.30	-0.10	Docile	Triumphant	0.69	0.57	0.63	Exuberant	Exuberant
					Self-satisfied	0.86	0.20	0.62	Exuberant	Exuberant
					Aggressive	0.41	0.63	0.62	Exuberant	Exuberant
Anger	-0.51	0.59	0.25	Hostile	Angry	-0.51	0.59	0.25	Hostile	Aggressive
irritation, fury, rage					Irritated	-0.58	0.40	0.01	Hostile	Aggressive
					Enraged	-0.44	0.72	0.32	Hostile	Aggressive
Hate	-0.60	0.60	0.30	Hostile	Hate	-0.56	0.59	0.13	Hostile	Aggressive
disgust, dislike, loathe					Disgusted	-0.60	0.35	0.11	Hostile	Aggressive
Disliking	-0.40	0.20	0.10	Hostile	Hostile but controlled	-0.24	0.42	0.09	Hostile	Aggressive
detest, hate, disgust					Disgusted	-0.60	0.35	0.11	Hostile	Aggressive
Fear	-0.64	0.60	-0.43	Anxious	Fearful	-0.64	0.60	-0.43	Anxious	Anxious
worried, scared					Upset	-0.63	0.30	-0.24	Anxious	Anxious
Shame	**-0.30**	**0.10**	**-0.60**	Anxious	Shamed	-0.57	0.01	-0.34	Anxious	Anxious
self-blame, embarrassed					Embarrassed	-0.46	0.54	-0.24	Anxious	Anxious

Remorse *penitent, self-anger*	**-0.30**	**0.10**	**-0.60**	Anxious	Guilty Regretful Repentant Sinful	-0.57 -0.52 -0.06 -0.30	0.28 0.02 0.06 0.22	-0.34 -0.21 -0.12 -0.01	Anxious Anxious Anxious Anxious	Anxious Anxious Anxious Anxious
Disappointment *despair, frustration*	-0.30	0.10	-0.40	Anxious	Despairing Frustrated	-0.72 -0.64	-0.16 0.52	-0.38 -0.35	Bored Anxious	Depressed Anxious
Reproach *contempt, appalled* *despise, disdain*	-0.30	-0.10	0.40	Disdainful	Reprehensible Contemptuous Scornful Disdainful	-0.09 -0.24 -0.35 -0.32	0.11 0.31 0.35 -0.11	0.06 0.21 0.29 0.05	Hostile Hostile Hostile Disdainful	Aggressive Aggressive Aggressive Disdainful
Fears-confirmed	-0.50	-0.30	-0.70	Bored	Dissatisfied Sad Discouraged	-0.50 -0.63 -0.61	0.05 -0.27 -0.15	0.13 -0.33 -0.29	Anxious Bored Bored	Anxious Depressed Depressed
Resentment *envy, jealousy*	-0.20	-0.30	-0.20	Bored	Frustrated Helpless Insecure	-0.64 -0.71 -0.57	0.52 0.42 0.14	-0.35 -0.51 -0.42	Anxious Anxious Anxious	Anxious Anxious Anxious
Distress *depressed, distressed* *displeased, dissatisfied*	**-0.40**	**-0.20**	**-0.50**	Bored	Distressed Depressed Displeased Dissatisfied	-0.61 -0.72 -0.55 -0.50	0.28 -0.29 0.16 0.05	-0.36 -0.41 -0.05 0.13	Anxious Bored Anxious Hostile	Anxious Depressed Anxious Hostile
Pity *compassion, sad-for,* *sorry-for*	**-0.40**	**-0.20**	**-0.50**	Bored	Sad Regretful	-0.63 -0.52	-0.27 0.02	-0.33 -0.21	Bored Anxious	Depressed Anxious

Table 6: Matching of the OCC emotions with emotions from the PAD theory

4.2 Changing the mood

One thing that becomes clear in table 6, is that which mood an emotion belongs to is not set in stone. The 'correct' mood is thus hard to define. In this chapter we are going to experiment with taking other moods for some emotions than we did in the previous chapter. We are going to do this by taking the orange colored emotions from table 6, and placing the emotion in one of the other moods it has according to the original PAD emotions and values. Then we will analyze whether this change of mood has influence on the patterns we found previously.

In table 7 we sorted all the OCC emotions according the mood they are in, as stated by the original PAD emotions and values. Where the mood was unclear, we took the mood where the majority of the emotions fell in. An example of this is *relief*, two emotions have the mood Docile and one Relaxed, so we picked Docile. When there was no majority, for example with *pity*, a mood different from the mood it was assigned in ALMA was chosen, to see the effect of the change. There is one exception: *fears-confirmed*. This emotion is hard to define since it has no tokens, and no clear match with the emotions from the PAD theory. We matched two emotions that fall in the mood Bored with it, and one that falls in the mood Anxious. Because we want to see what changes of mood do to the patterns we previously observed, we assigned the mood Anxious to *fears-confirmed*. We will return to this particular emotion and others that have no tokens and direct match with the emotions from the PAD theory later in this chapter. In table 7, one thing that catches the eye is the imbalance in emotions per mood: Exuberant and Anxious have six and even seven emotions, while Disdainful has zero. We will discuss this later in this chapter. For now, we can analyze the emotions and the logical structure that belongs to it in Steunebrink (2010) with this new arrangement, in the same way we did last chapter.

Mood	**Emotions**						
Exuberant (+P+A+D)	Joy	Pride	Gratification	Hope	Liking	Gloating	
Dependent (+P+A-D)	Happy-for	Love	Admiration	Gratitude			
Relaxed (+P-A+D)	Satisfaction						
Docile (+P-A-D)	Relief						
Hostile (-P+A+D)	Anger	Hate	Disliking	Reproach			
Anxious (-P+A-D)	Fear	Shame	Remorse	Resentment	Distress	Pity	Fears-confirmed
Disdainful (-P-A+D)							
Bored (-P-A-D)	Disappointment						

Table 7: Emotions sorted by their moods according to original PAD values

In figures 17 and 18, we indicated with colors the mood that an emotion belongs to with the new distribution according to the original PAD values. With the positively valenced reactions in figure 17, on first sight there are no major changes in the structure. With negatively valenced reactions in figure 18, we can see that

one mood disappears altogether (disdainful), and Bored has only one emotion. Furthermore, since the structure is an inheritance structure, the block of actions of agents does not lead to changes in mood anymore when going further down the inheritance structure. We analyzed the PAD values and their dependency on the P value in the same way as in the previous chapter, this can be seen in table 8. In this table, we can see that now more variables are independent of the P value; more than half of the emotions have a positive A value. This is desirable since it confirms the statement of Russell and Mehrabian (1977) that the three variables are nearly independent.

To add PAD values to the inheritance structure, we needed the PAD values not to change anymore when going down in the structure because of the compositionality. Visually we can see that this is the case for actions of agents. There are still problems with the lowest part of the structure, consequences of events which are actual and which have several other properties. There are also differences between the A and D values separately, we cannot say for instance these emotions always have a +A value. By using original PAD values, we saw that the variables become more independent. We still have not found a way to make the PAD theory and the OCC model completely compatible, but the integration of them has become more coherent. To see if there is a way to make the models compatible and integrate mood and emotion, we will look more closely at the emotion elicitation structure that we have used until now in the next section.

Emotions and eliciting condition	**+P**	**-P**	**+P**	**-P**
hope/fear: event - prospective	+A	+A	+D	-D
joy/distress: event - actual	+A	+A	+D	-D
satisfaction: event - actual + confirm desirable	-A	–	+D	–
fears-confirmed: event - actual + confirm undesirable	–	+A	–	-D
relief: event - actual + disconfirm undesirable	-A	–	-D	–
disappointment: event - actual + disconfirm desirable	–	-A	–	-D
happy-for/resentment: event - actual + presume desirable	+A	+A	-D	-D
gloating/pity: event - actual + presume undesirable	+A	+A	+D	-D
pride/shame: action - self	+A	+A	+D	-D
admiration/reproach: action - other	+A	+A	-D	+D
gratification/remorse: action - self + consequence	+A	+A	+D	-D
gratitude/anger: action - other + consequence	+A	+A	-D	+D
liking/disliking: object	+A	+A	+D	+D

Table 8: Emotions and their eliciting conditions with the separate PAD values based on the original PAD values from Russell and Mehrabian (1977)

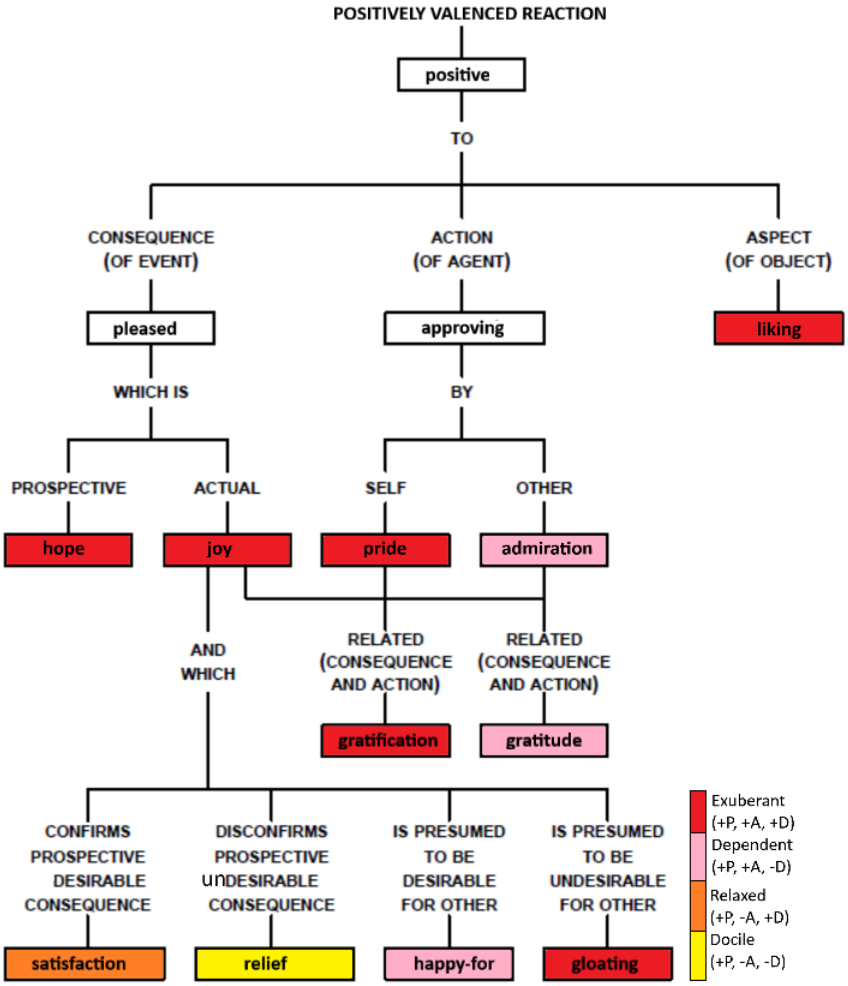

Figure 17: The logical structure of the emotion eliciting conditions with moods (+P) indicated with colors

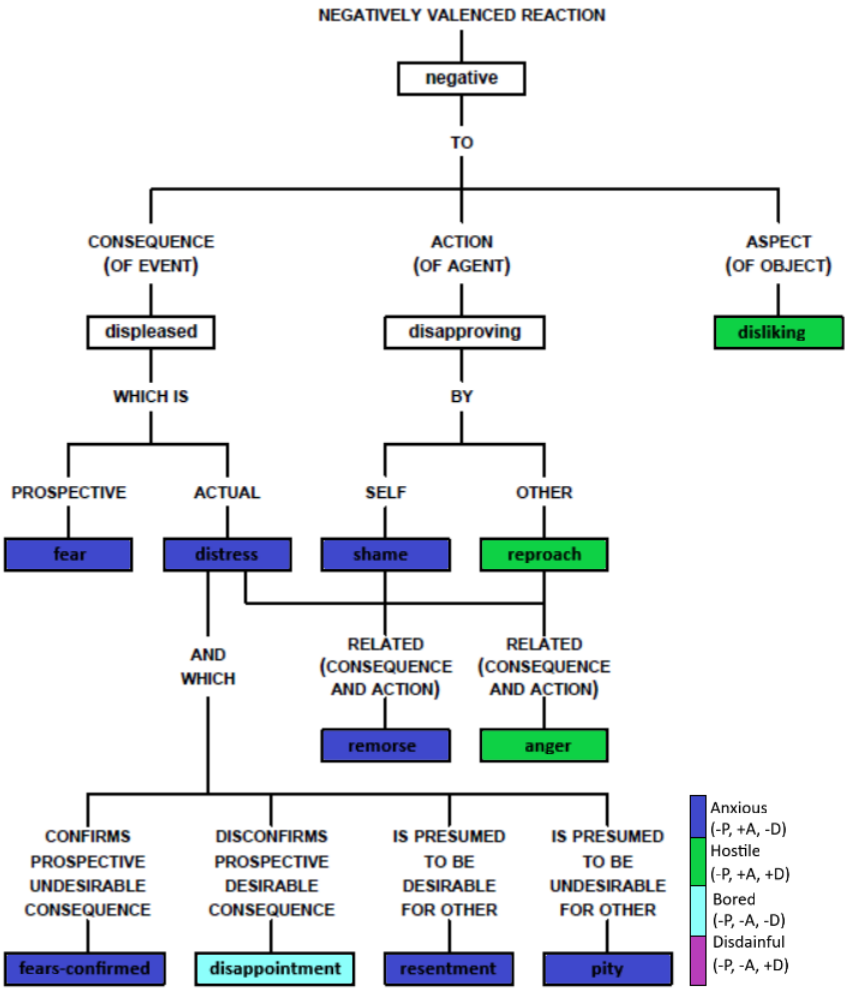

Figure 18: The logical structure of the emotion eliciting conditions with moods (-P) indicated with colors

4.3 The OCC structure vs the logical structure

Until now, we have used the structure of the OCC model that Steunebrink (2010) proposed. This structure differs from the structure Ortony et al. (1988) originally proposed. This is because for building a logical formalization, the structure has to be compositional, that is, complex expressions are established by its smaller components. The original structure is based on the focus of attention, and makes an earlier distinction based on this which causes it to not be compositional Steunebrink (2010). Even though the compositional structure of Steunebrink (2010) was confirmed to be accurate through personal communication with Ortony and Clore, we want to compare our findings with the original structure they proposed. This way we can maybe find out what it is about the emotions in the lower part of the structure that makes them currently unable to inherit the PAD values. In figure 19, the original structure is shown. The emotions are underlined with the color of the mood they belong to, according to the PAD values of ALMA. The main difference between this structure and the one used by Steunebrink (2010), is the part with 'consequences of events'. In the structure of Steunebrink (2010) (figure 17 and 18), *hope* and *fear* are not parents to other emotions. *Joy* and *distress* are here the parents of *fears-confirmed*, *disappointment*, *resentment*, *pity*, *satisfaction*, *relief*, *happy-for*, and *gloating*. In the original structure from OCC, *hope* and *fear* are the parents of *satisfaction*, *relief*, *fears-confirmed*, and *disappointment*. Those are the emotions that occur when a consequence is confirmed or disconfirmed. The other emotions, the ones that occur when a consequence has influence on someone else, do not have parents in the original structure. Otherwise the original structure is almost the same as the one Steunebrink (2010) uses, with only another difference in aspects of objects. In the original structure, liking and disliking leads to the emotions *love* and *hate*. In the structure of Steunebrink (2010) this step is not made, love and hate do not exist and the emotions there are *liking* and *disliking*.

When we look at which moods the emotions belong to according to the PAD values given in ALMA, we see roughly the same discrepancies occur as in the structure of Steunebrink (2010). The PAD values change when going lower in the inheritance structure at 'actions of agents' and at 'consequences for self - prospects relevant'. Figure 20 shows the mood the emotions are in according to the original PAD values, we have applied the original PAD values in the same way as was done on the structure of Steunebrink (2010) in the previous section. Just as in the structure of Steunebrink (2010), the problems around 'actions of agents' are now resolved. This indicates that the integration has become more coherent than before. However, there are still changes in the moods below *hope* and *fear*.

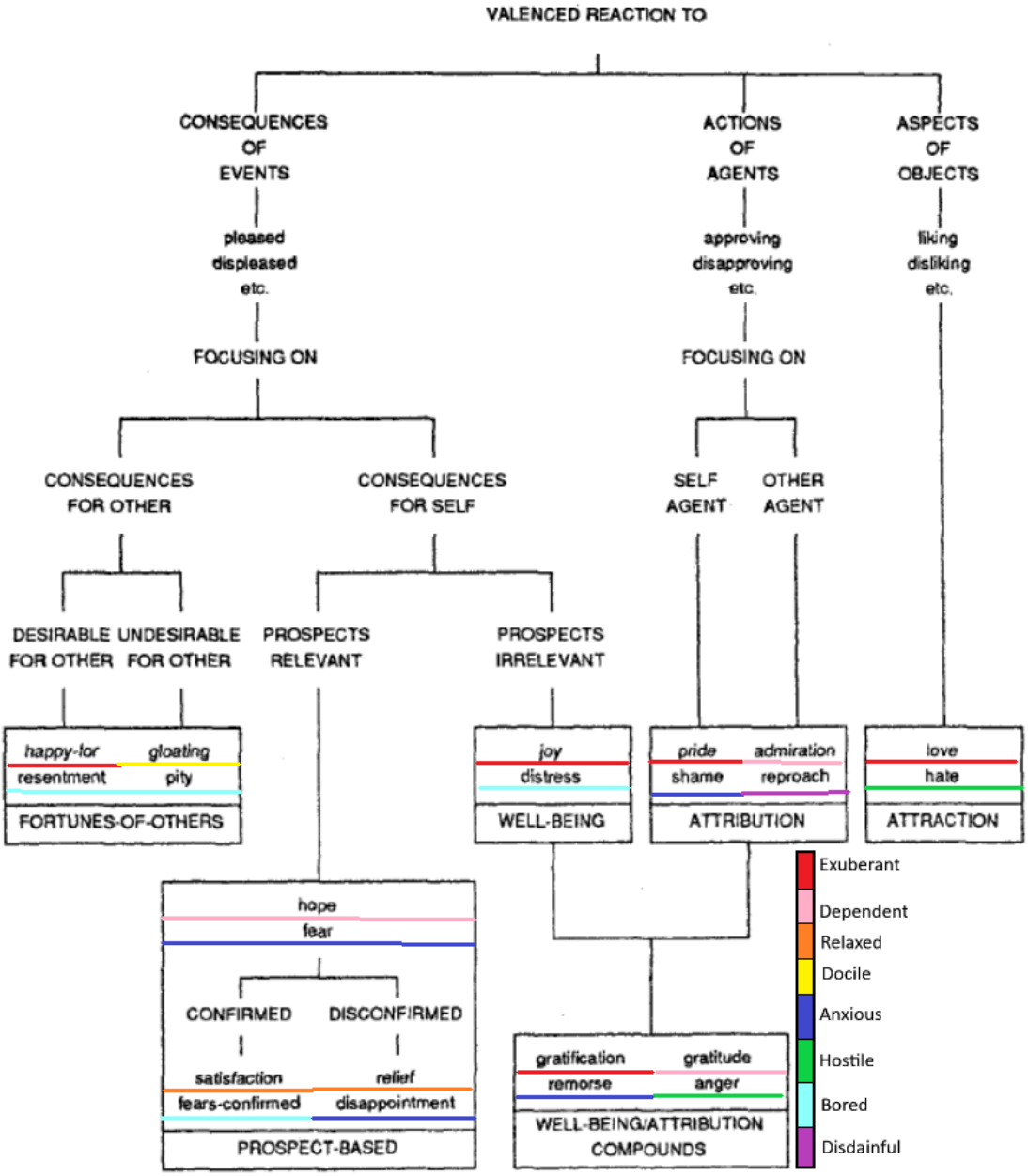

Figure 19: The original OCC structure with colors indication the mood per emotion. Moods are assigned with ALMA PAD values

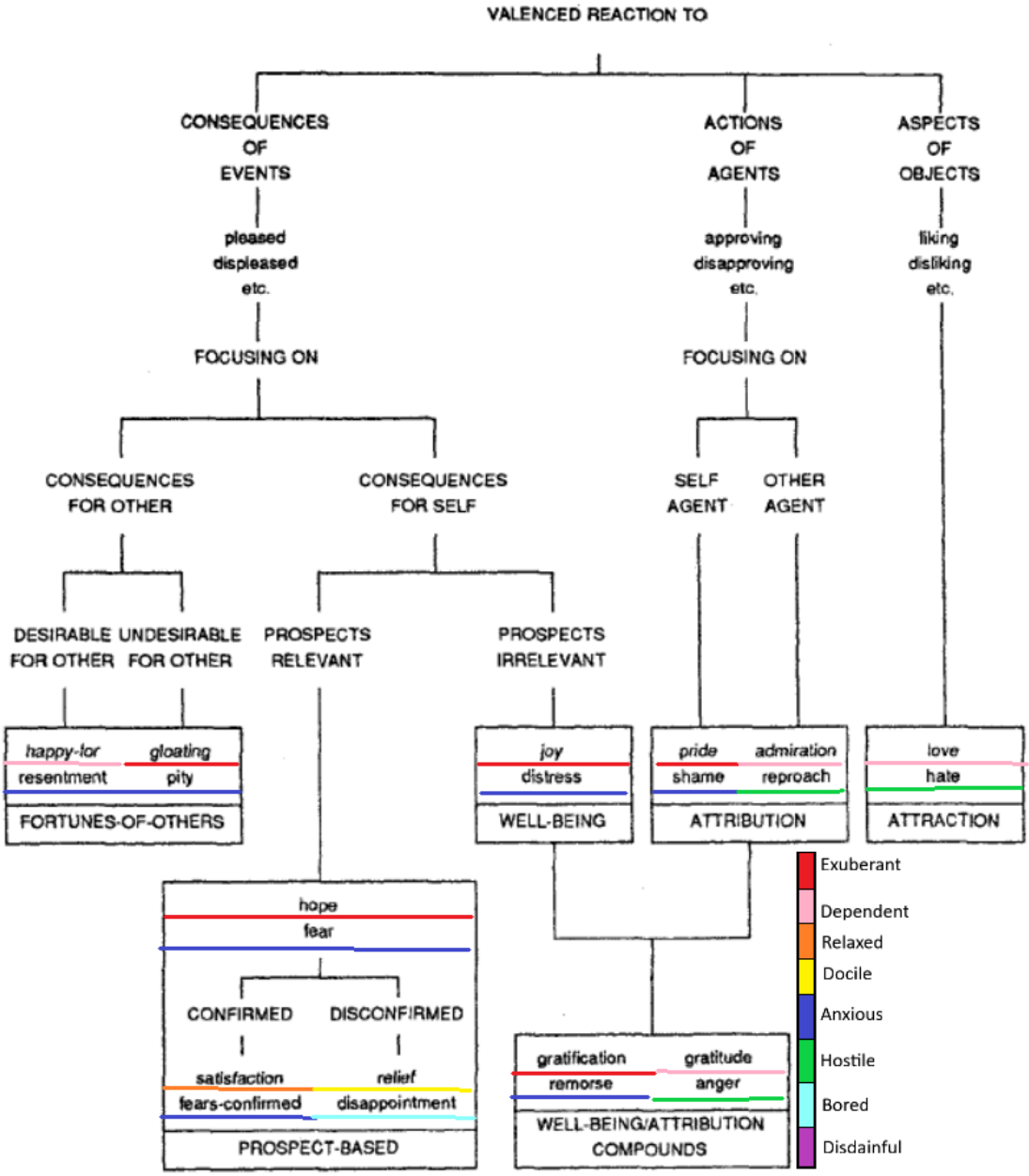

Figure 20: The original OCC structure with colors indication the mood per emotion. Moods are assigned with original PAD values

4.4 Intuitions

In the past sections we have critically studied the PAD values that each emotion has and therefore in which mood it falls. We tried to solve the problem that occurs when combining the inheritance structure of the OCC model with the PAD values: PAD values change lower in the inheritance structure, which defeats the purpose of compositionality and inheritance. When we studied the original PAD values found in experimental studies, we saw that the values became more independent. The balance between the OCC emotions became more off, some moods contained up to seven emotions while the smallest one contained zero. This might have been a reason for the ALMA model to assign other PAD values to some emotions. In ALMA it would have been more important to have at least one emotion for each mood octant, and preferably a more equal balance, since all the mood octants and emotions are needed for the computations. However, that does not justify that there are four emotions with exact the same PAD values, and that some emotions like *gloating* and *reproach* have completely different moods.

When looking at the original OCC structure, we can see the same problems occur in the inheritance structure as in the compositional structure of Steunebrink (2010). Here however, a distinction is made between the emotions that confirm or disconfirm a consequence, and emotions that are desirable or undesirable for others. The first group comes from *hope* and *fear*. Intuitively, this makes sense since *hope* and *fear* are prospective, and *satisfaction*, *relief*, *fears-confirmed*, and *disappointment* either confirm or disconfirm this prospect. In the compositional structure of Steunebrink (2010), these emotions are seen as 'actual' and therefore coming from *joy* and *distress*. This makes it easier to deal with in logic since the time stays the same, but it makes the lower level of the structure very detailed in descriptions and less elegant. Both structures have their downsides: the original structure is not compositional, the compositional structure of Steunebrink (2010) feels unbalanced at the bottom. Both structures cannot easily be combined in a logical way with PAD values.

To solve both problems at once, We can try to adjust the structure of the emotions. In figure 21 we propose a structure which is compositional, and which does justice to the emotions that have to do with prospects. Colors indicate the mood as was done in the previous figures: except that we use multiple colors here to indicate moods that are optional to the emotion. The main differences with the compositional structure of Steunebrink (2010) are the placement of the confirmed/disconfirmed consequences of prospective consequences, and the desirable or undesirable consequences for others. We made a distinction between those since we feel they are fundamentally different and come into existence with different emotions (*hope* and *fear* for the first and *joy* and *distress* for the second). This distinction also does justice to the diversity of these emotions. The new structure is compositional, and it also allows for PAD values to be incorporated in a more straightforward way than with the other structures.

When we look at the emotions under *joy* and *distress*, we see that there are multiple moods in which the emotions can fall. *Happy-for* can be either in Ex-

uberant or in Dependent according to the PAD values of both ALMA and the original values. *Gloating* can be in Docile according the the PAD values of ALMA, or in Exuberant with the original PAD values. *Resentment* can be in Bored with the PAD values of ALMA, or in Anxious with the original values. Pity can be in Bored or in Anxious. The emotions at the top of this category are less variable, *joy* is in Exuberant according to both theories, *distress* can be in Bored, Anxious or Hostile. We rule Hostile out here since only one of the tokens of this emotion falls in this category. Regardless of which moods are ultimately the correct ones for the emotions, we can manipulate them in such a way that the emotions inherit the PAD values from *joy* and *distress*. We can for instance say *happy-for* and *gloating* are in Exuberant, and *resentment* and *pity* are in Anxious or in Bored, depending on which mood we choose for *distress*. Of course we do not know the correct mood for these emotions, especially since there is no clear equivalent for the four emotions below *joy* and *distress*. But in this new structure, it is at least possible that it works with the PAD values. There are still four emotion left that now are in another place than previously: *satisfaction*, *relief*, *resentment*, and *pity*. These emotions are now below *hope* and *fear*. These are trickier to explain in terms of an inheritance structure: now all the PAD values can change, instead of just the A and the D value. For example, when *hope* (+P,+A,+D according to original PAD values) is disconfirmed, it leads to *disappointment* (-P, -A or +A, -D). This could have to do with the difference in time: When hoping something, you are either unsure of it or it is in the future. Either way, you have no current knowledge about the state this 'something' is in. When it is confirmed or disconfirmed, you gain this knowledge and therefore enter another state, a state where you do have knowledge about this event. In this case there would not be an inheritance relationship such as with the other emotions, it would be a change of state. Whether this would cause problems or not for formalizing this structure goes beyond the scope of this thesis, but it would definitely be fascinating to look at the possibilities of formalizing these ideas.

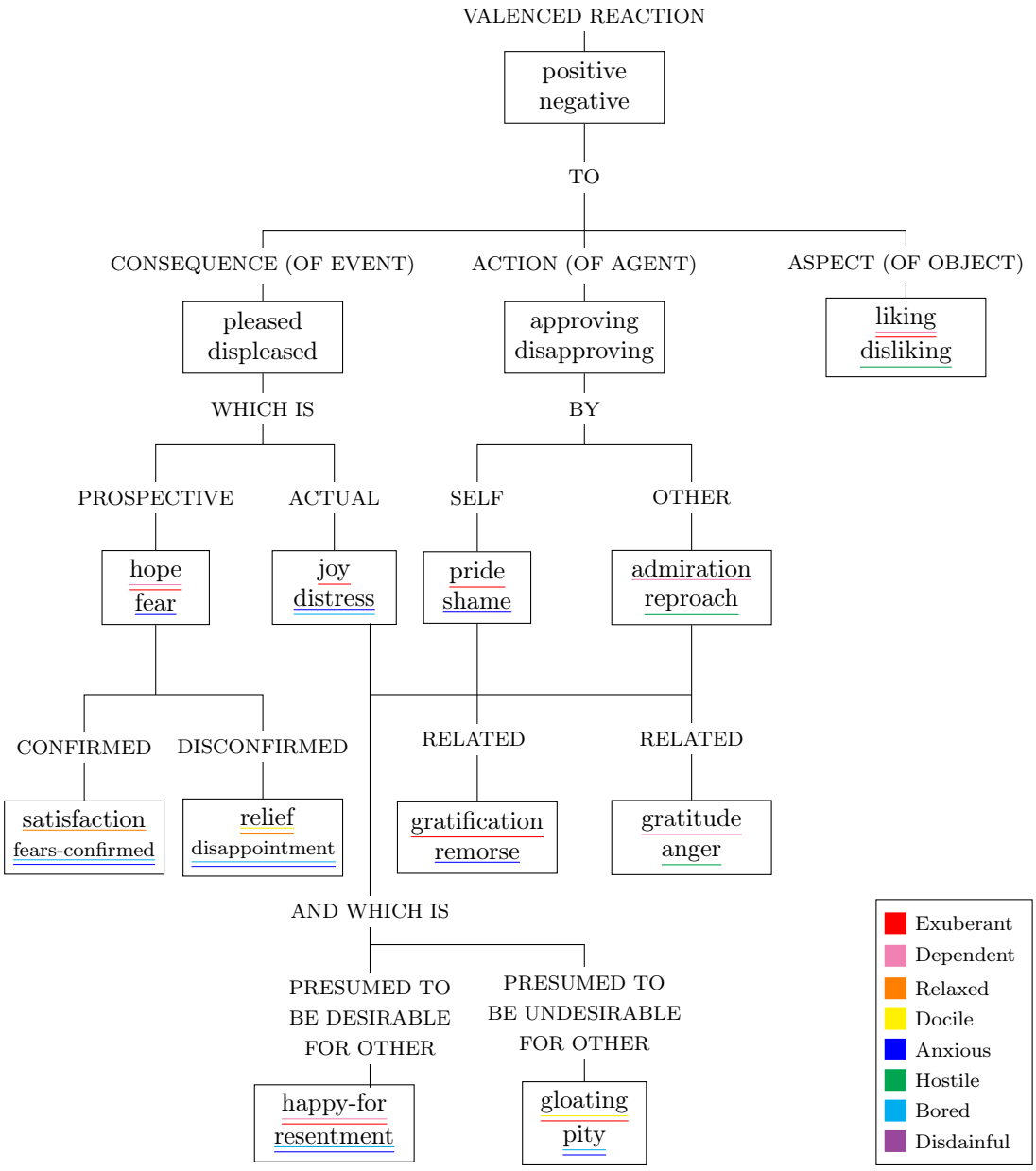

Figure 21: A proposal for an eliciting structure that is both compositional and compatible with moods and PAD values

5 Conclusion and discussion

In this thesis we aimed to integrate mood in models of agents with emotions. In order to do this we put forward four research questions; the main research question and three subquestions. They are repeated below.

- How can mood be integrated in models of agents with emotions?

1. What are the affective phenomena *mood* and *emotion* exactly, how do they differ, and what is their relation?
2. How are *emotion* and *mood* formalized in agent models?
3. How can *emotion* and *mood* be integrated?

In chapter two, we aimed to answer the first two subquestions by analyzing literature from psychology, computer science, and artificial intelligence. In the psychological literature, we discovered that a clear definition on *emotion* and *mood* was not easy. There is a general consensus on some aspects: emotions are shorter in duration than moods, and emotions are reactions to something that happens in the environment of the subject. Moods do not originate from a reaction but they emerge from someones mind. Even though an exact definition is not of high importance in our thesis, it helps to understand emotion and mood. The definition of emotion that we use in this thesis is the one from the OCC model: emotions are valenced reactions to events, agents or objects. For moods we did not find such a definition, since most theories were focused on the function of mood. Moods interact with emotions: moods are calculated from emotions and in turn influence new emotions and their intensity.

Emotions are formalized in several agent models such as ALMA, EMA, and the logical models of Steunebrink (2010) and Adam (2007). This is done in different ways: ALMA is based on appraisal theory and uses the PAD theory and the OCC model to compute emotions and moods, EMA is based on the psychological theory of Lazarus and therefore also adds another affective process to their model (coping). The logical model of Steunebrink (2010) is completely based on the OCC model, the logical model of Adam (2007) is similar to this one but with other logical assumptions and another formalization. ALMA was the only model that we studied that uses mood. Mood is formalized in ALMA using the PAD theory. Emotions are given a PAD value and then are used to calculate the mood. Mood then in turn influences new emotions.

We used this information about *mood* and *emotion* in chapter 3 to answer the last subquestion: How can emotion and mood be integrated? We especially wanted to formalize the concept of mood, since that has not been done before. We decided that the logical model of Steunebrink (2010) lent itself best for an extension with mood. We could get our main inspiration from how ALMA incorporated mood in their model. We discovered that adding this concept of mood as an extension would not be possible since the current structure of Steunebrink (2010) would not allow

for the moods and emotions to interact in the way they do in ALMA. Therefore we proposed a new structure that would explain the general flow and interaction of mood and emotion better. Now we could focus on what was needed to allow mood to be added to the logical model: adding PAD values to the emotions. In both ALMA and the logical model, emotions from the OCC model were used. Mood was defined by a different theory: the PAD theory. Therefore both theories had to be combined to let mood and emotion interact. ALMA proposed a mapping for this: They put PAD values on every emotion from the OCC model. We also used this mapping and started applying it to not only the emotions but also the logical structure Steunebrink (2010) provided. We hoped that we would find a pattern such that we could add the PAD values as predicates at certain places in the logic. This was not the case, therefore we separated the PAD values and tried again to find patterns. We found that there were few A and D values that were independent of the P value, and those who were, were not in a clear logical 'block' in the structure. These were interesting patterns, but would make it very hard to actually formalize mood in this way. Therefore in chapter 4 we take a closer look at the PAD values of the emotions, and the logical structure that we used. We explored several different options to check whether we can adjust the PAD values to make it easier to add them to emotions in the logic. We concluded that regardless of the value, there will be problems in the structure with changing values. Therefore we proposed a different structure that might allow for the PAD values being added in logical places, and for still having enough variance in the different moods. Whether this new structure is viable when formalizing it in logic remains to be seen.

With this information we can answer our main research question: How can mood be integrated with models of agents with emotions? In the logical model of Steunebrink (2010), we can integrate mood by changing the general structure on appraisal, experience and regulation, by adding PAD values to the emotions, and possibly by changing the structure of the emotion elicitation process. Integrating mood in a model of agents with emotions would be easier if the model in question was inspired by the PAD theory, because combining this theory with the OCC model leads to mainly structural difficulties. However, the OCC model gives a better groundwork for formalizing because of this structure. In the next section we will discuss questions that remain open, and downsides of our ideas.

5.1 Discussion and future work

Our findings from the previous chapters have several unanswered questions and lead to new questions. The first change we made was to the general structure of the emotion process: our proposed new structure was one that is a hybrid structure between the one from ALMA and the one from Steunebrink (2010) (figure 14). This structure solved our problem of where mood would be in the process, but leads to a new problem: the distinction between triggered and experienced emotions that is important in the OCC model is no longer there. This does not have to be bad, other computational models that are inspired by the OCC model also do not have

this distinction. When we do want this distinction, we would have to think of changing the structure such that there is a difference in the emotion eliciting phase and the emotion experience phase. This in turn would make it harder to integrate the process of mood. In this structure we also added new input data: personality. We made a suggestion of using the five-factor theory of McCrae and John (1992) but did not discuss it further. Depending on how it is added, it could be an important component in both the emotion elicitation and in the mood generation. In the structure we left out intensity by saying that the intensity of the emotions is determined by the percepts and concerns. The PAD theory has absolute values of the P, A, and D values and thereby determines the intensity of the emotions. By adding the PAD theory, even when not using exact values, it is hard to predict how intensities of emotions will behave. To do this we would have to find out how the intensity exactly occurs in the percepts and concerns, how mood can influence this intensity, how mood gets its intensity, what the relationship is between the PAD values and the intensity, and so on. We would also need to know more about the PAD values independently: is the intensity perfectly balanced between the three or is for example the A value more important for the intensity of an emotion or mood? These and related questions are fascinating topics for future work. In chapter 4 we went more into detail with the PAD values and the structures of the elicitation process. We analyzed the structures from the OCC model and from the logical model of Steunebrink (2010), and proposed a new structure of the elicitation process. This structure seemed to work better with PAD values and moods: it allowed for more variation in moods and possibly the inheritance structure could be intact with PAD values. Going into detail on this structure to find out whether it would also work with logic goes beyond the scope of this thesis but would definitely be interesting to see what this structure can give rise too. Lastly, we did not formalize any of our ideas in logic yet because of the obstacles we met. The first step was adding the PAD values in a way that would work with the elicitation structure of emotions, that is, lower in the structure, emotions should inherit its parents' values. With the new structure of this process that we proposed, it might be possible to achieve this while staying true to the PAD values that belong to the emotions (original or from ALMA). Even when the PAD values are all added and formalized in a proper way, we would still have to think about how to calculate a mood from them, and how to let mood influence the emotion elicitation process. Finally, we can think of many more interesting ideas regarding this topic, but we believe we made a first step to an integrated model of mood and emotion.

References

Adam, C. (2007). Emotions: from psychological theories to logical formalization and implementation in a bdi agent.

Adam, C., Herzig, A., and Longin, D. (2009). A logical formalization of the occ theory of emotions. *Synthese*, 168(2):201–248.

Bakker, I., van der Voordt, T., Vink, P., and de Boon, J. (2014). Pleasure, arousal, dominance: Mehrabian and russell revisited. *Current Psychology*, 33(3):405–421.

Bäuml, K.-H. and Kuhbandner, C. (2007). Remembering can cause forgetting—but not in negative moods. *Psychological Science*, 18(2):111–115.

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994a). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1):7–15.

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994b). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1-3):7–15.

Clore, G. L., Gasper, K., and Garvin, E. (2001). Affect as information. *Handbook of affect and social cognition*, pages 121–144.

Clore, G. L. and Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in cognitive sciences*, 11(9):393–399.

Dastani, M., Floor, C., and Meyer, J.-J. C. (2014). Programming agents with emotions. In *Emotion Modeling*, pages 57–75. Springer.

Dastani, M., Meyer, J.-J. C., et al. (2006). Programming agents with emotions. In *ECAI*, pages 215–219.

Ekman, P. E. and Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. Oxford University Press.

Evans, D. (2002). *Emotion: The science of sentiment*. Oxford University Press, USA.

Folkman, S. and Lazarus, R. S. (1988). The relationship between coping and emotion: Implications for theory and research. *Social science & medicine*, 26(3):309–317.

Frijda, N. H. (1987a). Emotion, cognitive structure, and action tendency. *Cognition and emotion*, 1(2):115–143.

Frijda, N. H. (1987b). *The emotions*. Cambridge University Press.

Frijda, N. H., Kuipers, P., and Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of personality and social psychology*, 57(2):212.

Gebhard, P. (2001). Enhancing embodied intelligent agents with affective user modelling. *User Modeling 2001*, pages 271–273.

Gebhard, P. (2005). Alma: a layered model of affect. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 29–36. ACM.

Gebhard, P., Kipp, M., Klesen, M., and Rist, T. (2003a). Adding the emotional dimension to scripting character dialogues. In *IVA*, pages 48–56. Springer.

Gebhard, P., Kipp, M., Klesen, M., and Rist, T. (2003b). Authoring scenes for adaptive, interactive performances. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 725–732. ACM.

Gebhard, P., Klesen, M., and Rist, T. (2004). Coloring multi-character conversations through the expression of emotions. In *Tutorial and Research Workshop on Affective Dialogue Systems*, pages 128–141. Springer.

Goleman, D. P. (1995). Emotional intelligence: Why it can matter more than iq for character, health and lifelong achievement.

Gratch, J. and Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306.

Gratch, J. and Marsella, S. (2014). Appraisal models. *The Oxford handbook of affective computing*, pages 54–67.

Gross, J. J. (2013). *Handbook of emotion regulation*. Guilford publications.

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

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American psychologist*, 46(8):819.

Lazarus, R. S. (1993). Coping theory and research: past, present, and future. *Psychosomatic medicine*, 55(3):234–247.

Lazarus, R. S. (2006). *Stress and emotion: A new synthesis*. Springer Publishing Company.

Marsella, S. and Gratch, J. (2006). *EMA: A computational model of appraisal dynamics*. na.

Marsella, S., Gratch, J., Petta, P., et al. (2010). Computational models of emotion. *A Blueprint for Affective Computing-A sourcebook and manual*, 11(1):21–46.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215.

Mehrabian, A. (1978). Measures of individual differences in temperament. *Educational and Psychological Measurement*, 38(4):1105–1117.

Mehrabian, A. (1980). *Basic Dimensions for a General Psychological Theory Implications for Personality, Social, Environmental, and Developmental Studies*.

Mehrabian, A. (1996a). Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, 48(2):86–92.

Mehrabian, A. (1996b). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4):261–292.

Mehrabian, A. and O'reilly, E. (1980). Analysis of personality measures in terms of basic dimensions of temperament. *Journal of Personality and Social Psychology*, 38(3):492.

Mehrabian, A. and Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.

Meyer, J.-J. C. (2004). Reasoning about emotional agents. In *Proceedings of the 16th European conference on artificial intelligence*, pages 129–133. IOS Press.

Meyer, J.-J. C. (2006). Reasoning about emotional agents. *International journal of intelligent systems*, 21(6):601–619.

Meyer, J.-J. C., van der Hoek, W., and van Linder, B. (1999). A logical approach to the dynamics of commitments. *Artificial Intelligence*, 113(1-2):1–40.

Moors, A., Ellsworth, P. C., Scherer, K. R., and Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, 5(2):119–124.

Nowlis, V. (1965). Research with the mood adjective check list.

Nowlis, V. (1970). Mood: Behavior and experience. In *Feelings and emotions: The Loyola symposium*, pages 261–277. Academic Press San Diego, CA.

Oatley, K. and Jenkins, J. M. (1996). *Understanding emotions*. Blackwell publishing.

Ortony, A., Clore, G. L., and Collins, A. (1988). *The cognitive structure of emotions*. Cambridge university press.

Picard, R. (1997). Does hal cry digital tears? emotions. and computers. *HAL's Legacy: 2001's Computer as Dream and Reality*, pages 279–304.

Reisenzein, R., Hudlicka, E., Dastani, M., Gratch, J., Hindriks, K., Lorini, E., and Meyer, J.-J. C. (2013). Computational modeling of emotion: Toward improving the inter-and intradisciplinary exchange. *IEEE Transactions on Affective Computing*, 4(3):246–266.

Russell, J. A. and Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, 11(3):273–294.

Salovey, P. and Mayer, J. D. (1990). Emotional intelligence. *Imagination, cognition and personality*, 9(3):185–211.

Scherer, K. R., Scherer, K. R., and Ekman, P. (1984). On the nature and function of emotion: A component process approach. *Approaches to emotion*, 2293:317.

Smith, C. A., Lazarus, R. S., et al. (1990). Emotion and adaptation. *Handbook of personality: Theory and research*, pages 609–637.

Steunebrink, B. R. (2010). *The logical structure of emotions*. PhD thesis, Utrecht University.

Steunebrink, B. R., Dastani, M., and Meyer, J.-J. C. (2009). The occ model revisited. In *Proc. of the 4th Workshop on Emotion and Computing*.

Steunebrink, B. R., Dastani, M., Meyer, J.-J. C., et al. (2007). A logic of emotions for intelligent agents. In *Proceedings of the National Conference on Artificial Intelligence*, volume 22, page 142. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236):433–460.

Velásquez, J. D. (1996). *Cathexis–a computational model for the generation of emotions and their influence in the behavior of autonomous agents*. PhD thesis, Massachusetts Institute of Technology.

Appendices

A The emotion specifications of the OCC model Ortony et al. (1988)

In this appendix the specifications of emotions from Ortony et al. (1988) are summarized. The intended reading of the specifications is described in section 2.2.3. The tokens are used in chapter 4 to find a better match between emotions from the OCC model and the PAD theory.

Joy Emotions
TYPE SPECIFICATION: (pleased about) a desirable event
TOKENS: contented, cheerful, delighted, ecstatic, elated, euphoric, feeling good, glad, happy, joyful, jubilant, pleasantly surprised, pleased, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event is desirable
EXAMPLE: The man was pleased when he realized he was to get a small inheritance from an unknown distant relative.

Distress Emotions
TYPE SPECIFICATION: (displeased about) an undesirable event
TOKENS: depressed, distressed, displeased, dissatisfied, distraught, feeling bad, feeling uncomfortable, grief, homesick, lonely, lovesick, miserable, regret, sad, shock, uneasy, unhappy, upset, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event is undesirable
EXAMPLE: The driver was upset about running out of gas on the freeway.

Happy-for emotions
TYPE SPECIFICATION: (pleased about) am evemt presumed to be desirable for someone else
TOKENS: delighted-for, happy-for, pleased-for, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the desirable event for the other is desirable for oneself
(2) The degree to which the event is presumed to be desirable for the other person
(3) The degree to which the other person deserved the event
(4) The degree to which the other person is liked
EXAMPLE: Fred was happy for his friend Mary because she won a thousand dollars.

Pity[3] emotions
TYPE SPECIFICATION: (Displeased about) an event presumed to be undesirable for someone else
TOKENS: compassion, pity, sad-for, sorry-for, sympathy, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the undesirable event for the other is undesirable for oneself
(2) The degree to which the event is presumed to be undesirable for the other person
(3) The degree to which the other person did not deserve the event
(4) The degree to which the other person is liked
EXAMPLE: Fred was sorry for his friend Mary because her husband was killed in a car crash.

TYPE SPECIFICATION: (Displeased about) an event presumed to be desirable for someone else
TOKENS: envy, jealousy, resentment, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the desirable event for the other is undesirable for oneself
(2) The degree to which the event is presumed to be desirable for the other person
(3) The degree to which the other person did not deserve the event
(4) The degree to which the other person is liked
EXAMPLE: The executive resented the large pay raise awarded to a colleague whom he considered incompetent.

Gloating Emotions
TYPE SPECIFICATION: (pleased about) an event presumed to be undesirable for someone else
TOKENS: Gloating, Schadenfreude, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the undesirable event for the other is desirable for oneself
(2) The degree to which the event is presumed to be undesirable for the other person
(3) The degree to which the other person deserved the event
(4) The degree to which the other person is liked
EXAMPLE: Political opponents of Richard Nixon gloated over his ignominious departure from office.

Hope emotions
TYPE SPECIFICATION: (pleased about) the prospect of a desirable event
TOKENS: anticipation, anticipatory excitement, excitement, expectancy, hope, hopeful, looking forward to, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the event is desirable
(2) The likelihood of the event
EXAMPLE: As he thought about the possibility of being asked to the dance, the boy was filled with hope.

Fear emotions
TYPE SPECIFICATION: (displeased about) the prospect of an undesirable event
TOKENS: apprehensive, anxious, cowering, dread, fear, fright, nervous, petrified, scared, terrified, timid, worried, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the event is undesirable
(2) The likelihood of the event
EXAMPLE: The employee, suspecting he was no longer needed, feared that he would be fired.

Satisfaction emotions
TYPE SPECIFICATION: (pleased about) the confirmation of the prospect of a desirable event
TOKENS: gratification, hopes-realized, satisfaction, etc.
VARIABLES AFFECTING INTENSITY:
(1) The intensity of the attendant hope emotion
(2) The effort expended in trying to attain the event
(3) The degree to which the event is realized
EXAMPLE: When he realized that he was indeed being asked to go to the dance by the girl of his dreams, the boy was gratified.

Fears-confirmed emotions
TYPE SPECIFICATION: (displeased about) the confirmation of the prospect of an undesirable event
TOKENS: fears-confirmed, worst fears realized
VARIABLES AFFECTING INTENSITY:
(1) The intensity of the attendant fear emotion
(2) The effort expended in trying to prevent the event
(3) The degree to which the event is realized
EXAMPLE: The employee's fears were confirmed when he learned that he was indeed going to be fired.

Relief emotions
TYPE SPECIFICATION: (pleased about) the disconfirmation of the prospect of an undesirable event
TOKENS: relief
VARIABLES AFFECTING INTENSITY:
(1) The intensity of the attendant fear emotion
(2) The effort expended in trying to prevent the event
(3) The degree to which the event is realized
EXAMPLE: The employee was relieved to learn that he was not going to be fired

Disappointment emotions
TYPE SPECIFICATION: (displeased about) the disconfirmation of the prospect of a desirable event
TOKENS: dashed-hopes, despair, disappointment, frustration, heartbroken, etc.
VARIABLES AFFECTING INTENSITY:
(1) The intensity of the attendant hope emotion
(2) The effort expended in trying to attain the event
(3) The degree to which the event is realized
EXAMPLE: The boy was disappointed when he realized that he would not be asked to the dance after all.

Pride emotions
TYPE SPECIFICATION: (approving of) one's own praiseworthy action
TOKENS: pride
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged praiseworthiness
(2) The strength of the cognitive unit with the actual agent
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: The woman was proud of saving the life of a drowning child

Shame emotions
TYPE SPECIFICATION: (disapproving of) one's own blameworthy action
TOKENS: embarrassment, feeling guilty, mortified, self-blame, self-condemnation, self-reproach, shame, psychologically uncomfortable, uneasy, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged blameworthiness
(2) The strength of the cognitive unit with the actual agent
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: The spy was ashamed of having betrayed his country

Admiration[4] **emotions**
TYPE SPECIFICATION: (approving of) someone else's praiseworthy action
TOKENS: admiration, appreciation, awe, esteem, respect, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged praiseworthiness
(2) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: The physicist's colleagues admired her for her Nobel-prize-winning work.

Reproach emotions
TYPE SPECIFICATION: (disapproving of) someone else's blameworthy action
TOKENS: appalled, contempt, despise, disdain, indignation, reproach, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged blameworthiness
(2) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: Many people despised the spy for having betrayed his country

Gratitude emotions
TYPE SPECIFICATION: (approving of) someone else's praiseworthy action and (being pleased about) the related desirable event
TOKENS: appreciation, gratitude, feeling indebted, thankful, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged praiseworthiness
(2) The degree to which the event is desirable
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: The man was grateful to the stranger for saving the life of his child

Anger emotions
TYPE SPECIFICATION: (disapproving of) someone else's blameworthy action and (being displeased about) the related undesirable event
TOKENS: anger, annoyance, exasperation, fury, incensed, indignation, irritation, livid, offended, outrage, rage, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged blameworthiness
(2) The degree to which the event is undesirable
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness EXAMPLE: The man was angry with his wife for forgetting to buy the groceries.

Gratification emotions
TYPE SPECIFICATION: (approving of) one's own praiseworthy action and (being pleased about) the related desirable event
TOKENS: gratification, pleased-with-oneself, self-satisfaction, smug, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged praiseworthiness
(2) The degree to which the event is desirable
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness
(4) The strength of the cognitive unit of the agent
EXAMPLE: The man was gratified by his son's achievements.

Remorse emotions
TYPE SPECIFICATION: (disapproving of) one's own blameworthy action and (being displeased about) the related undesirable event
TOKENS: penitent, remorse, self-anger, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree of judged blameworthiness
(2) The degree to which the event is undesirable
(3) deviations of the agent's action from person/role-based expectations (i.e., unexpectedness
(4)The strength of the cognitive unit with the agent
EXAMPLE: The spy felt remorse at the damage he had done in betraying his country.

Liking emotions
TYPE SPECIFICATION: (liking) an appealing object
TOKENS: adore, affection, attracted to, like, love, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the object is appealing
(2) The degree of familiarity with the object
EXAMPLE: John was filled with affection as he gazed at his newborn infant

Disliking emotions
TYPE SPECIFICATION: (disliking) an unappealing object
TOKENS: aversion, detest, disgust, dislike, hate, loathe, repelled-by, revulsion, etc.
VARIABLES AFFECTING INTENSITY:
(1) The degree to which the object is unappealing
(2) The degree of familiarity with the object
EXAMPLE: Mary disliked the concert so much that she left in the middle.

B Logic of Steunebrink (2010)

In this section we give a short overview of the logic of Steunebrink (2010) for the readers with a logical background. We will not explain the predicates and operators, for explanation and more details on the grounding of the logic see Steunebrink (2010).

B.1 Emotion Triggers

$Positive_i^T(X) \stackrel{\text{def}}{=} Perceive_i(X) \wedge Good_i(X)$
$Negative_i^T(X) \stackrel{\text{def}}{=} Perceive_i(X) \wedge Bad_i(X)$

$Perceive_i(X) \stackrel{\text{def}}{=} PerceiveConseq_i(X) \vee PerceiveAction_i(X) \vee PerceiveObject_i(X)$

$Good_i(X) \stackrel{\text{def}}{=} Des_i(X) \vee Praisew_i(X) \vee Appeal_i(X)$
$Bad_i(X) \stackrel{\text{def}}{=} Undes_i(X) \vee Blamew_i(X) \vee Unappeal_i(X)$

$Pleased_i^T(c) \stackrel{\text{def}}{=} PerceiveConseq_i(c) \wedge Des_i(c)$
$Displeased_i^T(c) \stackrel{\text{def}}{=} PerceiveConseq_i(c) \wedge Undes_i(c)$

$Approving_i^T(j:a) \stackrel{\text{def}}{=} PerceiveAction_i(j:a) \wedge Praisew_i(j:a)$
$Disapproving_i^T(j:a) \stackrel{\text{def}}{=} PerceiveAction_i(j:a) \wedge Blamew_i(j:a)$

$Liking_i^T(x) \stackrel{\text{def}}{=} PerceiveObject_i(x) \wedge Appeal_i(x)$
$Disliking_i^T(x) \stackrel{\text{def}}{=} PerceiveObject_i(x) \wedge Unappeal_i(x)$

$Hope_i^T(c) \stackrel{\text{def}}{=} Pleased_i^T(c) \wedge Prospective_i(c) Fear_i^T(c) \stackrel{\text{def}}{=} Displeased_i^T(c) \wedge Prospective_i(c)$

$Joy_i^T(c) \stackrel{\text{def}}{=} Pleased_i^T(c) \wedge Actual_i(c)$
$Distress_i^T(c) \stackrel{\text{def}}{=} Displeased_i^T(c) \wedge Actual_i(c)$

$Pride_i^T(j:a) \stackrel{\text{def}}{=} Approving_i^T(j:a) \wedge Cogunit_i(j)$
$Shame_i^T(j:a) \stackrel{\text{def}}{=} Disapproving_i^T(j:a) \wedge Cogunit_i(j:a)$
$Admiration_i^T(j:a) \stackrel{\text{def}}{=} Approving_i^T(j:a) \wedge \neg CogUnit_i(j:a)$
$Reproach_i^T(j:a) \stackrel{\text{def}}{=} Disapproving_i^T(j:a) \wedge \neg CogUnit_i(j:a)$

$Gratification_i^T(j:a,c) \stackrel{\text{def}}{=} PastPride_i^T(j:a) \wedge PastJoy_i^T(c) \wedge PerceiveRelated_i(j:a,c)$
$Remorse_i^T(j:a,c) \stackrel{\text{def}}{=} PastShame_i^T(j:a) \wedge PastDistress_i^T(c) \wedge PerceiveRelated(j:a,c)$
$Gratitude_i^T(j:a,c) \stackrel{\text{def}}{=} PastAdmiration_i^T(j:a) \wedge PastJoy_i^T(c) \wedge PerceiveRelated_i(j:a,c)$
$Anger_i^T(j:a,c) \stackrel{\text{def}}{=} PastReproach_i^T(j:a) \wedge PastDistress_i^T(c) \wedge PerceiveRelated_i(j:a,c)$

$Satisfaction_i^T(c,c') \stackrel{\text{def}}{=} Joy_i^T(c) \wedge PastHope_i^T(c') \wedge Confirms(c,c')$
$Fears-confirmed_i^T(c,c') \stackrel{\text{def}}{=} Distress_i^T(c) \wedge PastFear_i^T(c') \wedge Confirms(c,c')$
$Relief_i^T(c,c') \stackrel{\text{def}}{=} Joy_i^T(c) \wedge PastFear_i^T(c') \wedge Disconfirms(c,c')$
$Disappointment_i^T(c,c') \stackrel{\text{def}}{=} Distress_i^T(c) \wedge PastHope_i^T(c') \wedge Disconfirms(c,c')$

$Happy-for_i^T(c,j) \stackrel{\text{def}}{=} Joy_i^T(c) \wedge Presume_i Des_j(c)$
$Pity_i^T(c,j) \stackrel{\text{def}}{=} Distress_i^T(c) \wedge Presume_i Undes_j(c)$
$Gloating_i^T(c,j) \stackrel{\text{def}}{=} Joy_i^T(c) \wedge Presume_i Undes_j(c)$
$Resentment_i^T(c,j) \stackrel{\text{def}}{=} Distress_i^T(c) \wedge Presume_i Des_j(c)$

B.2 Basic Operators

$B_i\varphi$: Agent i believes φ (to be true)
$[i:a]\varphi$: After the execution of action a by agent i,φ holds.

Prevφ : In the previous state, φ was true.
Pastφ : Some time in the past, φ was true.
Futφ : Some time in the Future, φ was true.
$Fut^+\varphi \stackrel{\text{def}}{=} \neg\varphi \wedge Fut\varphi$

B.3 Events, actions, and objects

$Presume_i\varphi \stackrel{\text{def}}{=} B_i\varphi$
$New\varphi \stackrel{\text{def}}{=} \varphi \wedge \neg Prev\varphi$
$BelUpd_i(\varphi) \stackrel{\text{def}}{=} NewB_i\varphi$

$Prospective_i(\varphi) \stackrel{\text{def}}{=} FutUpd_i(\varphi) \vee UncUpd_i(\varphi)$
$Actual_i(\varphi) \stackrel{\text{def}}{=} BelUpd_i(\varphi)$

$FutUpd_i(\varphi) \stackrel{\text{def}}{=} NewB_iFut^+\varphi$
$UncUpd_i(\varphi) \stackrel{\text{def}}{=} New(\neg B_i\varphi \wedge \neg B_i\neg\varphi)$

$Done(i:a) \stackrel{\text{def}}{=} \langle i:a^-\rangle\top$
$PerceiveAction_i(j:a) \stackrel{\text{def}}{=} BelUpd_i(PastDone(j:a))$
$PerceiveRelated_i(j:a,\varphi) \stackrel{\text{def}}{=} BelUpd_i(Related(j:a,\varphi))$
$Related(i:a,\varphi) \stackrel{\text{def}}{=} Past(Done(i:a) \wedge New\varphi)$

$PerceiveObject_i(x) \stackrel{\text{def}}{=} BelUpd_i(object_x)$

B.4 Grounding

$M,s \vDash \langle i:\pi\rangle\varphi$ iff $\exists(M',s') \in S : M',s' \vDash \varphi$ and $((M,s),(M',s')) \in R_{i:\varpi}$
$M,s \vDash \text{Fut}\varphi$ iff $\exists(M',s') \in S : M',s' \vDash \varphi$ and $((M,s),(M',s')) \in (\bigcup R)*$
$M,s \vDash \text{Past}\varphi$ iff $\exists(M',s') \in S : M',s' \vDash \varphi$ and $((M',s'),(M,s)) \in (\bigcup R)*$
$M,s \vDash \text{Prev}\varphi$ iff $\exists(M',s') \in S : M',s' \vDash \varphi$ and $((M',s'),(M,s)) \in (\bigcup R)$

$M,s \vDash Des_i(\varphi)$ iff $\varphi \in Des(i)(M,s)$
$M,s \vDash Undes_i(\varphi)$ iff $\varphi \in Undes(i)(M,s)$
$M,s \vDash Praisew_i(j:a)$ iff $(j,a) \in Praisew(i)(M,s)$
$M,s \vDash Blamew_i(j:a)$ iff $(j,a) \in Blamew(i)(M,s)$
$M,s \vDash Appeal_i(x)$ iff $x \in Appeal(i)(M,s)$
$M,s \vDash Unappeal_i(x)$ iff $x \in Unappeal(i)(M,s)$
$M,s \vDash CogUnit_i(j)$ iff $j \in CogUnit(i)(M,s)$

$Confirms_i(\varphi,\psi) \stackrel{\text{def}}{=} B_i(\varphi \sqsubseteq \psi)$
$Disconfirms_i(\varphi,\psi) \stackrel{\text{def}}{=} B_i(\overline{\varphi} \sqsubseteq \psi)$

C Original PAD values found with emotions in Mehrabian and O'reilly (1980)

Term	N	Pleasure		Arousal		Dominance	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
1. Bold	27	.44*	.32	.61*	.24	.66*	.30
2. Useful	27	.70*	.20	.44*	.28	.47*	.40
3. Mighty	27	.48*	.37	.51*	.28	.69*	.31
4. Kind	27	.73*	.22	.19*	.32	.57*	.27
5. Self-satisfied	27	.86*	.10	.20	.40	.62*	.31
6. Admired	29	.81*	.21	.44*	.30	.51*	.34
7. Proud	29	.77*	.21	.38*	.34	.65*	.33
8. Interested	29	.64*	.20	.51*	.21	.17	.40
9. Arrogant	29	.00	.51	.34*	.44	.48*	.34
10. Inspired	29	.71*	.30	.63*	.21	.34*	.55
11. Excited	29	.62*	.25	.75*	.20	.38*	.29
12. Influential	28	.68*	.23	.40*	.33	.75*	.18
13. Aggressive	28	.41*	.30	.63*	.25	.62*	.24
14. Strong	28	.58*	.24	.48*	.30	.62*	.30
15. Dignified	28	.55*	.27	.22*	.40	.61*	.30
16. Powerful	28	.54*	.26	.45*	.36	.73*	.25
17. Elated	28	.50*	.47	.42*	.14	.23*	.36
18. Hopeful	29	.51*	.30	.23*	.33	.14	.41
19. Triumphant	29	.69*	.23	.57*	.19	.63*	.26
20. Joyful	29	.76*	.22	.48*	.26	.35*	.31
21. Capable	29	.70*	.24	.28*	.27	.61*	.31
22. Lucky	30	.71*	.19	.48*	.30	.37*	.29
23. Responsible	30	.35*	.29	.38*	.26	.49*	.37
24. Friendly	30	.69*	.23	.35*	.28	.30*	.27
25. Masterful	30	.58*	.25	.44*	.27	.69*	.25
26. Free	30	.81*	.14	.24*	.38	.46*	.33
27. Devoted	30	.49*	.25	.17*	.27	.10	.37
28. Domineering	27	.23*	.37	.40*	.21	.58*	.26
29. Aroused	30	.24*	.28	.57*	.26	.22*	.33
30. Concentrating	28	.42*	.25	.28*	.27	.39*	.31
31. Happy	29	.81*	.21	.51*	.26	.46*	.38
32. Egotistical	29	.24*	.34	.32*	.25	.50*	.31
33. Carefree	29	.78*	.21	.25*	.39	.41*	.31
34. Affectionate	29	.64*	.26	.35*	.34	.24*	.40
35. Vigorous	30	.58*	.22	.61*	.23	.49*	.24
36. Activated	30	.42*	.26	.58*	.21	.38*	.28
37. Alert	30	.49*	.25	.57*	.20	.45*	.26
38. Alone with responsibility	30	.33*	.34	.34*	.37	.48*	.36
39. Controlling	30	.47*	.26	.34*	.23	.66*	.25
40. Proud and lonely	27	.01	.43	.02	.32	.26*	.40
41. Enjoyment	30	.77*	.17	.44*	.26	.42*	.29
42. Serious	30	.27*	.22	.24*	.22	.42*	.27
43. Cooperative	31	.39*	.32	.13*	.27	.03	.34
44. Thankful	27	.61*	.25	.10	.34	−.13	.35

Term	N	Pleasure		Arousal		Dominance	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
45. Respectful	27	.38*	.39	.13	.29	−.08	.49
46. Appreciative	27	.55*	.26	.07	.30	−.14	.35
47. Loved	28	.87*	.17	.54*	.26	−.18*	.30
48. Grateful	28	.64*	.23	.16	.22	−.21*	.34
49. In love	28	.82*	.17	.65*	.24	−.05	.37
50. Anxious	28	.01	.45	.59*	.31	−.15	.32
51. Impressed	28	.41*	.26	.30*	.25	−.32*	.34
52. Surprised	29	.40*	.30	.67*	.27	−.13	.38
53. Sexually excited	29	.58*	.26	.62*	.20	−.01	.43
54. Wonder	30	.27*	.37	.24*	.35	−.17*	.26
55. Fascinated	29	.55*	.22	.51*	.23	−.07	.35
56. Awed	30	.18*	.34	.40*	.30	−.38*	.21
57. Overwhelmed	30	.14	.39	.45*	.34	−.24*	.32
58. Curious	28	.22*	.30	.62*	.20	−.01	.34
59. Relaxed	27	.68*	.30	−.46*	.38	.06	.49
60. Untroubled	27	.79*	.25	−.01	.46	.33*	.39
61. Modest	29	.27*	.34	−.06	.30	.12	.39
62. Secure	29	.74*	.22	−.13	.32	.03	.37
63. Nonchalant	30	.07	.35	−.25*	.31	.11	.33
64. Aloof	30	.16	.35	−.01	.30	.25*	.32
65. Leisurely	30	.58*	.35	−.32*	.33	.11	.33
66. Reserved	30	.01	.37	−.19*	.28	.02	.37
67. Protected	27	.60*	.35	−.22*	.37	−.42*	.40
68. Consoled	27	.29*	.50	−.19*	.33	−.28*	.38
69. Quiet	27	.19	.57	−.40*	.21	−.04	.39
70. Sheltered	29	.14	.53	−.36*	.32	−.44*	.24
71. Humble	28	.23*	.39	−.28*	.24	−.27*	.35
72. Solemn	29	.03	.39	−.32*	.26	−.11	.33
73. Reverent	28	.31*	.35	−.08	.26	−.29*	.34
74. Astonished	30	.16*	.26	.88*	.19	−.15*	.26
75. Disgusted	29	−.60*	.20	.35*	.41	.11	.34
76. Insolent	28	−.26*	.43	.21	.44	.20*	.31
77. Cruel	29	−.45*	.35	.48*	.32	.42*	.32
78. Irritated	29	−.58*	.16	.40*	.37	.01	.40
79. Defiant	28	−.16*	.30	.54*	.37	.32*	.42
80. Hate	28	−.56*	.16	.59*	.29	.13	.34
81. Hostile	29	−.42*	.31	.53*	.36	.30*	.32
82. Angry	29	−.51*	.20	.59*	.33	.25*	.39
83. Mildly annoyed	29	−.28*	.16	.17*	.28	.04	.31
84. Enraged	29	−.44*	.25	.72*	.29	.32*	.44
85. Contempt	29	−.23*	.39	.31*	.33	.18*	.29
86. Selfish	29	−.34*	.31	.09	.33	.31*	.43
87. Reprehensible	24	−.09	.36	.11	.24	.06	.32
88. Contemptuous	25	−.24*	.26	.31*	.31	.21*	.32
89. Scornful	28	−.35*	.21	.35*	.27	.29*	.32
90. Suspicious	29	−.25*	.23	.42*	.21	.11	.32
91. Skeptical	29	−.22*	.28	.21*	.25	.03	.33

Term	N	Pleasure		Arousal		Dominance	
		Mean	SD	Mean	SD	Mean	SD
92. Burdened with responsibility	30	−.08	.41	.28*	.31	.19	.47
93. Cold anger	30	−.43*	.29	.67*	.27	.34*	.44
94. Hostile but controlled	30	−.24*	.31	.42*	.35	.09	.38
95. Crushed	27	−.69*	.30	.03	.47	−.50*	.23
96. Frustrated	27	−.64*	.18	.52*	.37	−.35*	.30
97. Distressed	27	−.61*	.17	.28*	.46	−.36*	.21
98. Insecure	27	−.57*	.34	.14	.42	−.42*	.29
99. Humiliated	27	−.63*	.18	.43*	.34	−.38*	.30
100. Hungry	29	−.44*	.26	.14	.33	−.21*	.27
101. Fearful	29	−.64*	.20	.60*	.32	−.43*	.30
102. Terrified	29	−.62*	.20	.82*	.25	−.43*	.34
103. Embattled	27	−.37*	.24	.40*	.42	−.02	.39
104. Helpless	29	−.71*	.18	.42*	.45	−.51*	.32
105. Troubled	28	−.63*	.16	.16	.40	−.40*	.24
106. Startled	28	−.09	.34	.65*	.29	−.33*	.32
107. Anguished	29	−.50*	.30	.08	.46	−.20*	.34
108. Shamed	29	−.57*	.23	.01	.46	−.34*	.30
109. Displeased	29	−.55*	.21	.16	.34	−.05	.41
110. Embarrassed	29	−.46*	.30	.54*	.26	−.24*	.40
111. Upset	30	−.63*	.20	.30*	.39	−.24*	.24
112. Defeated	30	−.61*	.24	.06	.39	−.32*	.23
113. Pain	29	−.58*	.21	.41*	.31	−.34*	.28
114. Quietly indignant	26	−.28*	.35	.04	.36	−.16	.40
115. Repentant	25	−.06	.55	.06	.32	−.12	.43
116. Sinful	28	−.30*	.36	.22*	.33	−.01	.41
117. Shy	29	−.15	.33	.06	.30	−.34*	.28
118. Guilty	29	−.57*	.19	.28*	.38	−.34*	.28
119. Weary with responsibility	30	−.27*	.34	.02	.28	−.01	.35
120. Angry but detached	30	−.42*	.22	.28*	.41	−.03	.33
121. Confused	30	−.53*	.20	.27*	.29	−.32*	.28
122. Dissatisfied	30	−.50*	.22	.05	.28	.13	.32
123. Regretful	30	−.52*	.24	.02	.32	−.21*	.28
124. Tense	30	−.33*	.35	.58*	.32	−.11	.39
125. Disdainful	25	−.32*	.32	−.11	.27	.05	.29
126. Depressed	27	−.72*	.21	−.29*	.44	−.41*	.28
127. Despairing	27	−.72*	.21	−.16	.34	−.38*	.25
128. Lonely	29	−.66*	.35	−.43*	.36	−.32*	.30
129. Meek	29	−.19	.58	−.25*	.32	−.41*	.42
130. Burdened	29	−.66*	.21	−.03	.41	−.26*	.36
131. Timid	28	−.15	.41	−.12	.37	−.47*	.31
132. Bored	28	−.65*	.19	−.62*	.24	−.33*	.21
133. Feeble	29	−.42*	.35	−.20	.42	−.46*	.31
134. Nauseated	29	−.61*	.25	−.01	.28	−.36*	.33
135. Inhibited	30	−.54*	.26	−.04	.40	−.41*	.23
136. Fatigued	30	−.18	.48	−.57*	.29	−.29*	.29
137. Rejected	29	−.62*	.24	−.01	.38	−.33*	.27
138. Subdued	29	−.17	.35	−.26*	.25	−.18*	.34

Term	N	Pleasure		Arousal		Dominance	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
139. Impotent	30	−.53*	.35	−.13	.40	−.29*	.32
140. Ennui	16	−.45*	.39	−.43*	.36	−.17	.42
141. Blasé	26	−.29*	.33	−.51*	.25	−.16*	.30
142. Haughty and lonely	27	−.47*	.28	−.24*	.34	−.13	.26
143. Listless	27	−.45*	.41	−.59*	.28	−.24*	.28
144. Deactivated	28	−.46*	.32	−.43*	.41	−.46*	.27
145. Weary	29	−.18	.38	−.33*	.42	−.24*	.35
146. Snobbish and lonely	30	−.62*	.26	−.19*	.34	−.14*	.28
147. Uninterested	30	−.47*	.26	−.50*	.22	−.08	.24
148. Detached	30	−.37*	.34	−.26*	.28	−.14*	.29
149. Discontented	30	−.53*	.19	−.16	.41	−.26*	.30
150. Discouraged	30	−.61*	.25	−.15	.32	−.29*	.32
151. Sad	30	−.63*	.23	−.27*	.34	−.33*	.22