



Universiteit  
Leiden  
The Netherlands

## **Affect and Learning: a computational analysis**

Broekens, D.J.

### **Citation**

Broekens, D. J. (2007, December 18). *Affect and Learning: a computational analysis*. Leiden Institute of Advanced Computer Science (LIACS), Faculty of Science, Leiden University. Retrieved from <https://hdl.handle.net/1887/12537>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12537>

**Note:** To cite this publication please use the final published version (if applicable).

7

# Affect and Formal Models

*Formalizing Cognitive Appraisal Theory*



In this chapter we take a theoretical approach towards computational modeling of emotion. Affect in this chapter is thus interpreted in a broader sense, as in *related to emotion*. This is different from the interpretations of affect presented in Chapter 2 to 6. To avoid any potential misunderstanding, in this chapter we use the term emotion, not affect. We present a formal way in which emotion theories can be described and compared with the computational models based upon them. We apply this formal notation to cognitive appraisal theory, a family of cognitive theories of emotion, and show how the formal notation can help to advance appraisal theory and help to evaluate computational models based on cognitive appraisal theory: the main contributions of this chapter. Although this chapter is quite different from the others, it fits within the general approach: that is, the use of computational models to evaluate emotion theories. As such it can be viewed as a high-level analysis of issues associated with computational modeling of emotion.

Cognitive appraisal theories (CATs) explain human emotions as a result of the subjective evaluation of events that occur in the environment. Recently, arguments have been put forward that discuss the need for formal descriptions in order to further advance the field of cognitive appraisal theory. Formal descriptions can provide detailed predictions and help to integrate different CATs by providing clear identification of the differences and similarities between theories. A computational model of emotion that is based on a CAT also needs formal descriptions specifying the theory on which it is based. In this chapter we propose a formal notation for the declarative semantics of the structure of appraisal. We claim that this formalism facilitates both integration of appraisal theories as well as the design and evaluation of computational models of emotion based on an appraisal theory. To support these claims we show how our formalism can be used in both ways: first we integrate two appraisal theories; second, we use this formal integrated model as basis for a computational model after identifying what declarative information is missing in the formal model. Finally, we embed the computational model in an emotional agent, and show how the formal specification helps to evaluate the computational model.

## 7.1 Introduction

Computational models of emotion are used in a wide variety of artificial emotional agents. In general, such a model is based on a cognitive appraisal

theory (CAT) (note that the model of affect and affective feedback we have used in Chapter 2 to 5 are not based on cognitive appraisal theory). CATs explain human emotions as a result of the subjective evaluation of events. However, such theories typically lack the necessary detail to base a computational model upon (Gratch & Marcella, 2004). As a result, it is difficult to evaluate if the computational model correctly implements the theory.

Further, to advance the field of appraisal theory, it is essential that cognitive appraisal theories can be integrated and compared with each other. Thus, building computational models of emotion *and* advancing the field of appraisal theory are in need of a representation of appraisal theory that enables systematic analysis. This is the focus of our chapter.

More specific, we propose a formalism to describe the structure of appraisal. That is, we propose a formal notation for the behavior of processes that play a role in appraising a situation, how these processes are linked to each other, what the resulting emotions could be, etc. In this chapter we show that different cognitive appraisal theories can be described using the same formal notation, that such formal representations can be used to compare and integrate CATs and that the formal representation can be used to systematically analyze computational models of emotion.

Such formal description of a specific CAT can be used, for example, to prove that the happy expression on the face of a child, that just noticed it arrived at a large rollercoaster park with extremely exciting rollercoasters and a couple of flags, must be due to an appraisal of the situation that involves the expectancy of intrinsic pleasantness. If I would have a robot, the formalism can be used in approximately the same way. While developing the robot, I would use the formalism to understand why it shows a certain emotion. Assuming a specific CAT, the formalism can be used to decide whether its artificial emotion of fear is potentially correct after I have proposed to go to a rollercoaster park. At first, I might be tempted to start to debug the robot, but the formal description of the CAT on which its emotions are based can show me that its emotion might be genuine as it potentially results from a negative appraisal of the rain (reflecting its fear to rust).

This informal introduction gives some intuition for the need and use of formal representations of appraisal theory. In this chapter we propose a formalism to describe the structure of appraisal (Section 7.3) and we elaborate on two ways in which this formalism can be used: (1) we use it to integrate two different appraisal theories (Section 7.4), and (2) we use it to analyze a computational model of emotion we developed (Section 7.5). Before continuing the main line of this

chapter, we first give a cognitive definition of emotion, some more detail on the development and use of artificial emotional agents, and a more detailed description of the problem we address.

### 7.1.1 Emotion

In cognitive psychology, emotion is often defined as a psychological state or process that functions in the management of goals, needs, desires and concerns of an individual (we refer to these four terms as *goals*). This state consists of physiological changes, feelings, expressive behaviors, cognitive activity and inclinations to act (e.g., Roseman & Smith, 2001). Emotion is elicited by the evaluation of an event in relation to the accomplishment of the agent's goals. Thus, an emotion is a heuristic that relates events to the agent's goals (Oatley, 1999). Additionally, emotions are used in non-verbal communication.

### 7.1.2 Artificial Emotions

Inspired by this heuristic and communicative aspect of emotion, computational models of emotion are embedded in a variety of intelligent agents. The development of artificial emotional agents is useful, and can be applied to a wide variety of domains. These domains include electronic tutors (Heylen et al., 2003), human-robot interaction (Breazeal, 2001; Chapter 5), virtual agents in VR training environments (Henninger et al., 2002), agents targeted at decision-making and planning (Coddington & Luck, 2003) and adaptive agents that use emotion or affect to control learning parameters (Belavkin, 2004; Chapter 3-4). For example, research shows that a robot's emotional expression influences human caretaking behavior (Breazeal, 2001), of which the following is a nice anecdote. When human subjects interacted with Kismet (the emotional robot) and Kismet reacted sad or distressed to the actions of the human, the subjects were visibly distressed and looked questioning to the researchers as if they wanted to say "am I doing something wrong?" A second example is a recent study by Partala and Surakka (2004) that shows that affective intervention in human-computer interaction has a positive effect on the human, both emotionally as well as in terms of the subject's problem solving performance. Positive words resulted in smiling as well as better problem solving performance.

### 7.1.3 Cognitive Appraisal Theory

The majority of computational models of emotion embedded into intelligent agents are based on cognitive appraisal theory. Such theories of emotion attempt to explain why a certain event results in one emotional response rather than

another and why a certain emotion can be elicited by different events. The key concept of most CATs is that the subjective cognitive evaluation of events in relation to the agent's goals is responsible for emotion (Roseman & Smith, 2001). More generically one can say that events have to be evaluated as having personal meaning or relevance (van Reekum, 2000). This evaluation is called *appraisal*. It is generally accepted that physiological changes and other non-cognitive factors can influence the actual appraisal of events. Although previously most appraisal theories assumed that appraisal was a necessary and sufficient condition for emotion (Roseman & Smith, 2001), currently it is seen as an important component of emotion.

#### 7.1.4 How to Interpret Artificial Emotions in Relation to a CAT?

The “brain” of artificial intelligent agents is often based on a belief-desire-intention (BDI) architecture (Jennings, Sycara & Wooldridge, 1998). If cognitive evaluation of events in relation to the agent's goals is sufficient for emotion, then the addition of such an evaluation of events related to the beliefs, desires and intentions of an artificial agent is sufficient for computational emotions. This partly explains the current popularity of appraisal theories as basis for emotional agents.

However, appraisal theories are currently described in a way that is insufficiently precise as a specification for a computational model of emotion (Gratch & Marsella, 2004). As a result, many computational models are inspired by structural theories of appraisal—i.e., theories that describe the structural relations between events, appraisal processes and emotions—and implemented using artificial intelligence mechanisms. During implementation, designers are forced to make many assumptions about the exact mechanisms of appraisal. This results in a large gap between the structural theory of appraisal and the resulting computational model of emotion.

In addition to this, artificial agents have a more and more complex design. These agents are approaching a point at which inspection of the agent's program and internal state is no longer efficient to “debug” the agent’s design. We predict that in the future it will no longer be feasible to try to understand an agent's unexpected behavior by purely investigating its inner workings. Instead, a formal investigation of its behavior will be a necessary component of this process of understanding (Broekens & DeGroot, 2006), just like we need to ask a person about why he/she does something instead of *only* looking at neuroimaging data.

### 7.1.5 Advancing Appraisal Theory Needs Comparison and Integration

Apart from the problem of using appraisal theories as basis for computational models, another problem—directly related to appraisal theory—exists. Although most appraisal theories share the assumption that cognitive appraisal is an important part of emotion, many different appraisal theories exist (Reisenzein, 2001; Frijda & Mesquita, 2000; Smith and Kirby, 2000; Scherer 2001). Comparison between, and convergence of these theories is difficult, but important in order to advance the field of appraisal theory. Formalization of structural theories of appraisal can help to solve these problems in two different ways. First, formal descriptions facilitate comparison, convergence and integration of theories, because assumptions and relations between concepts are clarified (Wehrle & Scherer, 2001). Second, computational modeling of emotion is a powerful way of analyzing appraisal theories in a formal way (Wehrle & Scherer, 2001). Formal descriptions facilitate the evaluation of computational models, thereby contributing to the analysis of appraisal theories.

### 7.1.6 Aim and Scope of This Chapter

The main contribution of this chapter is an abstract-level, theory independent, set-based formalism that can be used to describe the structure of appraisal as describe by a cognitive appraisal theory. This formalism addresses the two issues introduced above.

- First, how can we advance cognitive appraisal theory? We argue that our formalism facilitates comparison and integration of CATs. We use our formalism as a tool to integrate the Stimulus Evaluation Check theory (Scherer, 2001) and Appraisal Detector Model (Smith & Kirby, 2000), two prominent and recent CATs. Our formalism can be used to describe the behavior of the processes involved in appraisal. It does not address the issue of how to formally describe and reason about what a certain emotion *is* in terms of *specific* beliefs, desires and intentions of a BDI agent (e.g., Meyer, 2004).
- Second, how to formally specify a structural appraisal theory, so that the resulting formal description can be used as basis for the specification and evaluation of the emotional behavior of an artificial agent? We argue that our formalism narrows the gap between appraisal theory and computational model, and we show how such a formal specification can be used as basis for a computational model of emotion we have developed. We also show how this specification helps to evaluate the computational model.



The structure of this chapter is as follows. First, we introduce the relation between computational models, structural theories of appraisal and process theories of appraisal. Then we introduce the actual formalism in Section 7.3. While the introduction of Section 7.3 is essential for understanding the rest of the chapter, the parts that detail the formalism are recommended to the mathematically oriented reader. Less mathematically oriented readers will find Section 7.4—showing how the formalism can be used as a tool to facilitate the integration of appraisal theories—as well as Section 7.5—demonstrating how a formal description of a structural model of appraisal can be used as basis for a computational model—more interesting. Section 7.6 discusses issues around formalization, and related approaches.

## 7.2 Appraisal Theory: Structure, Process and Computation

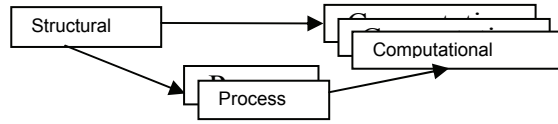
A common classification of CATs is based on a structural versus a process-based description (Roseman & Smith, 2001). Structural theories of appraisal (also called black-box models or structural models) describe the structural relations between:

- the environment of an agent and perception of this environment: *perception*;
- the agent's appraisal processes that interpret the perceived environment in terms of values on a set of subjective measures, called appraisal dimensions. An appraisal dimension influences emotion and can be considered as a variable—e.g., agency or valence—, used to express the result of the appraisal of a perceived object—e.g., a friend. This process of evaluation is called *appraisal*;
- the processes that relate these values to the agent's emotions: *mediation*.

Process theories of appraisal describe, in detail, the cognitive operations, mechanisms and dynamics by which the appraisals, as described by the structural theory, are made and how appraisal processes interact (Reisenzein, 2001). In other words, a structural theory of appraisal aims at describing the declarative semantics of appraisal, while a process theory of appraisal complements this description with procedural semantics. In this chapter we adopt the terms *structural model* and *process model* respectively, and use *appraisal theory/model* when referring to cognitive theories/models of appraisal in general.

A computational model is a model that is composed of operations that unambiguously control the behavior of a device. These operations may use available input data. If there is a sequence of such operations that maps a specific input to a specific behavior (output), an *algorithm* is said to exist for that mapping. The devices are essentially serial, but parallel execution can be either simulated in one such device using threading, or effectuated using multiple

communicating devices. In this chapter, we define a *computational model* as a structured collection of interacting algorithms that operate serially or in parallel, with operations that are eventually reducible to the Turing machine level.



**Figure 7.1.** Three possible mappings between structural, process and computational models of emotion.

In Broekens and DeGroot (2006) we have analyzed the relation between cognitive appraisal theory and computation. We have argued that it is useful to have a theory-independent formal notation to describe structural appraisal theories (i.e., the behavior of processes that play a role in appraising a situation, how these processes are linked to each other, what the resulting emotions could be, etc.). For clarity, we summarize the conclusion here.

In general, there is a generic-to-specific relation between structural, process and computational models. Structural models are the basis of computational- and process models, and process models are also the basis of computational models. In this case "basis of" usually means that a model *A* that is the basis of a model *B* contains less details than model *A*, and therefore different model *B* instantiations are possible based on model *A* (Figure 7.1). Although this is true in general, in Broekens and DeGroot (2006) we have argued that the difference between a structural, process and computational description is also one of kind, not just of different degrees of detail; all three models are equally important for cognitive appraisal theory. We have also shown that a formal description of the structural model is needed for the following reasons:

- to advance appraisal theory. A formal description facilitates comparison, convergence and integration of appraisal theories, and the process of formalization helps theory refinement;
- to build computational models of emotion based on structural theories of appraisal. First, process models of appraisals should coexist with computational models, not take their place. Second, before designing computational models at the algorithmic level, declarative information is needed on the processes that are responsible for perception, appraisal and mediation as defined by the appraisal theory. Third, objective information is needed to evaluate the consistency between computational model and appraisal theory, and reuse of this information seems very useful. We need a declarative description of the processes that are responsible for an agent's emotion, in order to evaluate if the agent's *unexpected* emotion resulting from an experimental situation is due to a problem in the agent's architecture, or

due to a mismatch between our interpretation of the situation and the agent's interpretation.

Typically, a common formal notation should enable formal description of a structural model such that this description includes the following data (of which many are also relevant to process models; Reisenzein, 2001):

- What is the nature and level (van Reekum, 2000) of processes; deliberative, automatic, innate?
- What is the relation between (results of) perception and appraisal processes.
- When and how are these processes activated? Are there thresholds? Can activation be sub-threshold?
- What kind of input and output (representations) a certain process needs/produces?
- Does a process continuously output results or periodically (how often)?
- How many and what perception, appraisal and mediating processes exist?
- Is information activation binary or gradual? E.g., how strongly must a certain event be perceived for it to be input for a certain appraisal process?
- What is the number of different appraisal dimensions, their activation range and the responsible processes?

### 7.3 A Set-Based Formalism for the Structure of Appraisal

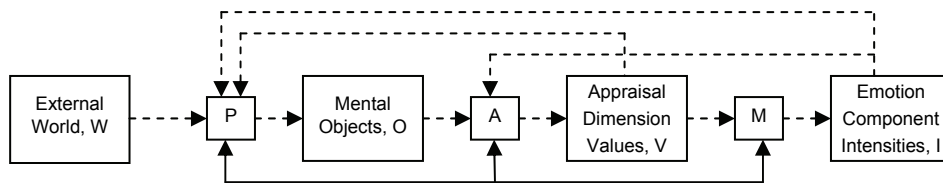
In this section we introduce the basic concepts of the formalism we propose to describe structural theories of appraisal. Later sections explain its use in some detail. Our formalism is set-based and built around sets of perception processes, appraisal processes and mediating processes (Figure 7.2). The notation used for these three types of processes and the accompanying terminology are borrowed from Reisenzein (2001). The external world,  $W$ , is the set of all events and objects that can respectively occur and reside in the environment. Perception processes, the set  $P$ , filter, select and translate information from the external world, and produce *mental objects*—representations of the external world suitable for appraisal. We define the set of mental objects produced by the perception processes, the set  $O$ , as the current content of working memory. Appraisal processes, the set  $A$ , evaluate the mental objects produced by the perception processes and assign a combination of appraisal dimension values, the set  $V$ , to these objects. Mediating processes relate appraisal information to emotions. Thus, mediating processes, the set  $M$ , relate appraisal dimension values to emotion-component intensities, the set  $I$ .

Perception processes also perceive the agent's current appraisal dimension values and current emotion components. These two kinds of information are

translated to mental objects. Since in our formalism only perception processes can put information in working memory, the emotion-component intensities,  $I$ , and appraisal information,  $V$ , must be perceived before the agent is able to use these two kinds of information in appraisal. This is consistent with the idea that appraisal is a cognitive evaluation of perceived objects in working memory. Additionally, the separation between cognitive emotional information—i.e.,  $V$  and  $I$  perceived by  $P$ —and non-cognitive emotional information— $I$  influencing  $A$ —enables the specification of appraisal processes that are biased by a specific combination of emotional feedback (i.e., none, non-cognitive, cognitive, or both). This enables, for example, explicit specification of appraisal structures involved in coping, re-appraisal and strategic use of emotions. This ability is important for the completeness of our formalism.

To describe the *structural relations* between elements in the sets of perception, appraisal and mediating processes, our formalism allows the specification of process dependencies. For example, some process dependencies can be defined as excitatory relations, while others can be defined as inhibitory relations between processes.

The concepts of the formalism are detailed in the rest of this section. To facilitate understanding of the formalism, we demonstrate its use by showing how the static (hypothetical) appraisal structure of a baby can be defined. The baby can be exposed to a barking dog or its mother, resulting in different emotions.



**Figure 7.2.** Graphical overview of the assumed structure of appraisal underlying our formalism. Dotted arrows denote potential inputs for processes, while normal arrows denote potential process dependencies. The external world contains events that can be perceived. Perception processes perceive event, appraisals and emotion-component intensities and map these to mental representations (including beliefs, goals, etc.). Appraisal processes appraise these representations in the context of the current emotion-component intensities, by mapping them to appraisal dimension values (e.g., an object is moderately arousing and moderately goal conducive), which are again mapped to emotion-component intensities by mediating processes (e.g., the current set of appraisals results in a smile and a feeling of excitement). For details see text.

### 7.3.1 World, Perception Processes and Objects of Appraisal

**Definition 7.1.1:**  $W=\{w_1,\dots,w_n\}$  is the set of all observable objects and events in the environment of the agent<sup>1</sup>.

**Definition 7.1.2:**  $O\subseteq PO$  is the current content of working memory, assuming that  $PO=\{po_1,\dots,po_n\}$  is the set of all potential mental objects with  $po_i=(t, any\_object\_name)$ ,  $po_i\in PO$  and  $t\in OT$ ,  $OT$  being the set of mental object types as defined in Definition 7.1.3.

**Definition 7.1.3:**  $OT=\{t_1,\dots,t_n\}$  is the set of type names—(O)bject (T)ypes—used to specify mental object types (e.g. belief, desire, goal, plan, etc.).

**Definition 7.1.4:** If we define  $V$  as the set of appraisal dimension values (see Definition 7.2.2) and  $I$  as the set of emotional-response-component intensities (see Definition 7.3.2) then  $P=\{p_1,\dots,p_n\}$  is the set of all perception processes available to the agent, with  $p_i:P(W\cup V\cup I)\rightarrow P(PO)$ ,  $p_i\in P$  such that  $\forall o\in O \exists p\in P \exists x\in P(W\cup V\cup I)$  with  $o\in p(x)$ . In words, a perception process  $p_i$  typically maps a portion of the agent's environment, several of the agent's current appraisal dimension values and several of its emotional-response-component intensities to one or more mental objects. These objects are the ones that can be in working memory<sup>2</sup>. Thus, we assume that if an object is in working memory then there must be a perception process producing it.

In our baby example the baby's world initially contains two objects: *mom* and *dog*, represented by two distinct noise levels  $m$  and  $d$ ,  $W=\{m, d\}$ . The baby can perceive these objects with her only perception function called hear,  $p_h$ , that perceives noise levels  $m$  and  $d$ .  $P=\{p_h\}$ , with  $p_h(\{m\})=\{po_m\}$ ,  $p_h(\{d\})=\{po_d\}$ ,  $p_h(\{m, d\})=\{po_m, po_d\}$  and for all other inputs  $x$ ,  $p_h(x)=\emptyset$ . Thus  $p_h$  maps  $m$  and  $d$  to mental objects  $PO=\{po_m, po_d\}$ . The set  $OT$  contains one element,  $OT=\{belief\}$ , thus  $po_m=(belief, mom)$  and  $po_d=(belief, dog)$ . The baby has two beliefs, *mom* is here and the *dog* is here. The set  $O$  is empty; we thus assume that the baby has not perceived anything.

---

<sup>1</sup> Note that we use  $n$  as a finite but arbitrary number to denote multiple elements in a set. When  $i$  is used as element index, we mean for any  $1\leq i\leq n$ . Two sets both with  $n$  elements do not necessarily have the same number of elements. When they do, another subscript is used, e.g.,  $m$ . Also,  $P(S)$  is used to denote the powerset of set  $S$ .

<sup>2</sup> Note that different perception processes could perceive the same object at the same time, even if they use different information. For example, an agent both smells and sees a person.

### 7.3.2 Appraisal Processes, Appraisal Dimensions and Values

**Definition 7.2.1:**  $D=\{d_1,\dots,d_n\}$  is the set of appraisal dimensions, containing elements like suddenness and pleasantness.

**Definition 7.2.2:**  $V=\{v_1,\dots,v_n\}$  is the set of current appraisal dimension values with  $v_i=(o, d, r)$ ,  $v_i \in V$ , and  $o \in O$ ,  $d \in D$  and  $r \in [-1,1]$ . In words,  $v_i$  is a tuple of a one-dimensional appraisal result attributed to one mental object, or,  $v_i$  is the result of appraising an object in terms of one appraisal dimension.

**Definition 7.2.3:**  $A=\{a_1,\dots,a_n\}$  with  $a_i:P(O \cup I) \rightarrow P(V)$ ,  $a_i \in A$  such that  $\forall v \in V \exists a \in A \exists x \in P(O \cup I)$  with  $v \in a(x)$ . Again in words,  $a_i$  is an appraisal process that interprets mental objects in the context of emotional-response-component intensities and attributes appraisal dimension values to other mental objects<sup>3</sup>. Appraisal can be biased by the current emotion, explaining  $I$  in the powerset of the input for the appraisal processes. Also, some appraisal processes may be relevant to emotion only through their relation with other appraisal processes. In this case these “indirect” appraisal processes assign only zero values to evaluated mental objects.

To continue our baby example, the baby has two appraisal processes, *pleasure* and *arousal*. Both assign tuples of values  $[-1,1]$  and appraisal dimensions to mental objects. There are two appraisal dimensions with almost the same name as the appraisal processes. Thus  $A=\{a_p, a_a\}$  and  $D=\{pleas, arous\}$ . The dog produces noise, so the baby appraises the dog as arousing and unpleasant. So,  $a_p(\{po_d\})=\{(po_d, pleas, -0.5)\}$  and  $a_a(\{po_d\})=\{(po_d, arous, 0.5)\}$ . For all other inputs  $x$ ,  $a_p(x)=\emptyset$  and  $a_a(x)=\emptyset$ . The set  $V$  currently is empty, as  $O$  is empty. Here, we ignore the formal description of the soothing voice of the baby’s mother, as such things tend to defy all attempts at formalization.

### 7.3.3 Formalizing the Mediating Processes

**Definition 7.3.1:**  $E=\{e_1,\dots,e_n\}$  is the set of possible components of the emotional response, like certain subjective feelings, facial expressions, physiological reactions and action tendencies.

---

<sup>3</sup> Note that the mental objects to which an appraisal value is attributed are not necessarily the same as the objects used in the appraisal process. Also note that the introduction of appraisal value  $r$  introduces a problem if different appraisal processes produce a result on the same appraisal dimension. For example, if two appraisal processes produce the same (*object, dimension, value*) tuple then only one is in the set  $V$  (per the definition of sets). However, this could mean that the total intensity of the appraisal dimension is invalid. Since the appraisal value is from the set of real numbers  $[-1, 1]$  we assume that this never happens, as it is always possible to pick a real number close enough to another one but different.

**Definition 7.3.2:**  $I = \{i_1, \dots, i_n\}$  is the set of emotional-response-component intensities with  $i_i = (e, r)$ ,  $i_i \in I$ ,  $r \in [-1, 1]$  and  $e \in E$  (note that we slightly overload notation here by using subscript  $i$  with variable  $i$ ). In words,  $i_i$  is the intensity of one specific emotional-response component, e.g., a heart rate of 0.5 (on some scale). Appraisal theories typically assume that appraisal dimension values, not emotional-response-component intensities, are attributed to objects. This explains the lack of a mental object in  $i_i$ .

**Definition 7.3.3:**  $M = \{m_1, \dots, m_n\}$  with  $m_i: P(V) \rightarrow P(I)$ ,  $m_i \in M$  such that  $\forall i \in I \exists m \in M \exists x \in P(V)$  with  $i \in m(x)$ . In words,  $m_i$  produces emotional-response-component intensities based on appraisal dimension values<sup>4</sup>. Note that the definitions of  $m_i$  and  $a_i$  follow a common appraisal conception that appraisals are directed at objects, but emotions can be objectless.

Our baby has three emotions: *calm*, *distressed* and *neutral*,  $E = \{calm, dis, neut\}$ . The baby has one mediating process  $M = \{m_e\}$  that relates  $V$  (the set of assigned appraisal dimension values) to  $I$  (the set of emotion component intensities) in the following way:  $m_e(\{(o_d, pleas, -0.5), (o_d, arous, 0.5)\}) = \{(calm, 0), (dis, 0.5), (neut, 0)\}$ . For all other inputs  $x$ ,  $m_e(x) = \emptyset$ . This means that if and only if the baby appraises a situation as arousing and negative, the resulting emotion is distress with intensity 0.5. Again,  $I$  is empty as  $V$  is empty; we assume the baby is currently not appraising something.

### 7.3.4 Dependency between Processes

Our formalism represents processes connected to each other via different kinds of guarded dependencies. To be able to define the notation for dependency relations between processes, we first define guards and dependency types.

**Definition 4.1:** The set  $G = \{g_1, \dots, g_n\}$  of guards is a set of second-order predicates over the elements of the sets  $P$ ,  $O$ ,  $A$ ,  $D$ ,  $V$ ,  $M$ ,  $E$  and  $I$ , and over the variable  $r$ , being the actual value of elements in the set  $V$  and the intensity of the emotional response components of the set  $I$ . This allows the definition of conditional dependencies between processes.

**Definition 4.2:** The set  $LT = \{n_1, \dots, n_n\}$  is a set of dependency type names—(L)ink (T)ypes—used to identify the nature of the dependency between two processes (e.g., inhibitory, causal, correlation, information flow, parallelism, etc.). Again we slightly overload notation by using  $n_n$ .

---

<sup>4</sup> Same as previous note but for elements in  $M$ .

**Definition 4.3:** Let  $L$  be the set  $L=\{l_1, \dots, l_n\}$  with  $L \subseteq PP \times PP \times G \times N$  and  $PP = P \cup A \cup M$ . The elements of  $L$  define dependencies—(L)inks—between processes constrained by the following. For a tuple  $(p, q, g, n)$ , with  $p, q \in PP$ ,  $g \in G$  and  $n \in N$ , processing in  $q$  is influenced in the way described by  $n$  only if the guard  $g$  is true and process  $p$  is active itself<sup>5</sup>.

For our baby, there are four dependencies  $L=\{l_1, l_2, l_3, l_4\}$  between the perception, appraisal and emotion generation processes. These dependencies define a causal activation relation:

- $l_1=(p_h, a_p, (\exists x x \in O), activation)$ ,
- $l_2=(p_h, a_a, (\exists x x \in O), activation)$ ,
- $l_3=(a_a, m_e, (\exists x x \in V \wedge x=(d, i) \wedge i < 0), activation), d \in D$
- $l_4=(a_p, m_e, (\exists x x \in V \wedge x=(d, i) \wedge i < 0), activation), d \in D$

These dependencies thus define that if and only if the baby hears something (has perceived an object, i.e.,  $\exists x x \in O$ ) the appraisal processes must be activated, after which mediating processes are again activated.

### 7.3.5 Data Constraints

The activation conditions of processes can be defined using the above mentioned dependencies and guards. To allow the specification of data constraints that must hold according to the theory, we define a set  $H$  of constraints, again containing second-order predicates. For example, if an appraisal intensity greater than 0.5 for the novelty dimension exists, there must be an emotional-response-component intensity greater than 0 for the orientation response. These constraints also allow formalization of what should happen when there are two appraisal values for the same appraisal dimension, e.g., the baby hears a large and a small dog, both appraised as arousing resulting in two appraisal values loading on the same appraisal variable. Now a data constraint can be used to specify that both values should be, e.g., added. These data constraints are global, and not attached to process dependencies, like the guards used to represent activation conditions.

---

<sup>5</sup> Note that when a structural theory only mentions the type of the dependencies between processes without mentioning any activation conditions,  $G$  can be defined as  $G=\{true\}$ , so that all dependencies have a guard that is always true and only the type of dependency is used. Second, although we could extend the formal notation by allowing multiple guards or types per dependency, this does not add expressive power to the notation itself since the sets  $N$  and  $G$  can be filled by an arbitrary number of conjunctions. When actually using the formalism to describe an appraisal theory, multiple guards and types per dependency are definitely allowed to simplify the resulting description of the model.



**Definition 7.5.1:** The set  $H=\{h_1, \dots, h_n\}$  of guards is a set of second-order predicates over the elements of sets  $P, O, A, D, V, M, E$  and  $I$ , and over the variable  $r$ , being the actual value of elements in the set  $V$  and the intensity of the emotional response components of the set  $I$ .

### 7.3.6 What Does the Baby Example Tell Us?

We have formally described the “structural theory” for our baby’s hypothetical appraisal structure. For example, if we see the baby crying, we can prove that the baby must be appraising the situation as arousing and unpleasant. We can thus use the formal description to analyze structural relations between emotion processes of our baby. Now imagine a baby (or agent) with a much more complex appraisal structure. If we see it crying while we are trying to make cuddling noises, we might be surprised about this unexpected reaction. However, the formal appraisal structure could be used to, e.g., investigate an alternative possibility: our cuddling noises are appraised as unpleasant and arousing. This would mean, e.g., that the formal model predicts high skin conductance and increase in heartbeat. This is a verifiable hypothesis, and can now be tested. In short, we can use the formal description to evaluate, in a systematic way, whether an emotion is expected or not according to a certain structural theory.

Now, imagine that our theory actually cannot explain why the baby cries (e.g., because skin conductance is predicted to be high but is low in reality), and that a second theory exists that can. We can now formally compare these theories and make explicit the differences between both, so that we are able to explain why the second correctly explains the baby’s crying. The sets of processes and dependencies of one theory can be systematically compared with those of another. This is a much more verifiable, understandable and repeatable process than comparing textual representations of structural theories. Comparing theories using our formal notation is the topic of the next section.

## 7.4 Using Formal Notation to Compare and Integrate Cognitive Appraisal Theories

To show that the formal notation presented above can be used as a tool to compare and integrate different appraisal theories, we present a more serious example than our hypothetical baby. We use our formalism to integrate Scherer's (2001) Stimulus Evaluation Checks (SEC) model and Smith and Kirby's (2000) Appraisal Detector Model (ADM) process model. We call this model the SSK model (Scherer, Smith and Kirby). Our goal is to show the utility of formal notations in the domain of emotion theory and the power of our proposed notation

in particular. We do not argue that the model we present in this section is the best integration of both theories. For the same reason we have limited ourselves to parts of both theories, the model we present here is not to be interpreted as a complete integration of all aspects of both theories.

#### 7.4.1 Scherer's SEC Model

This model is based on the idea that appraisal processes evaluate stimuli in a certain sequence (for simplicity, in this chapter stimulus and event are assumed to be the same). Five different types of appraisal processes exist related to the evaluation of novelty, pleasantness, goal/need conduciveness, coping potential and norm/self compatibility. These appraisal processes exist at three levels, the sensory-motor level, the schematic level and the conceptual level. Appraisal processes take different forms depending on the level they operate on. An overview of these forms is given in Table 7.1. For the current integration we restrict ourselves by excluding norm/self compatibility.

	<b>Novelty</b>	<b>Pleasantness</b>	<b>Goal/Need conduciveness</b>	<b>Coping potential</b>
<b>Sensory-Motor level (innate)</b>	Sudden, intense stimulation	Innate preferences/aversions	Basic needs	Available energy
<b>Schematic level (automatic)</b>	Familiarity: schema matching	Learned preferences or aversions	Acquired needs, motives	Body schema (automated knowledge of what the body can do, how it functions, etc.)
<b>Conceptual level (deliberative)</b>	Expectations: cause/effect, probability	Recalled, anticipated, or derived positive-negative estimates	Conscious goals, plans	Problem-solving ability

**Table 7.1.** Overview of the stimulus checks related to novelty, pleasantness and coping potential existing at the sensory-motor, schematic and conceptual level (Scherer, 2001).

In general, *sensory-motor* level appraisal processes are related to biological needs and drives and to biological mechanisms, and are mostly genetically determined. *Schematic-level* appraisals are based on learned knowledge organized into schemas. *Conceptual level* appraisal processes are based on propositional-symbolic, cortical mechanisms that require consciousness (Scherer, 2001). Higher levels are used to appraise the situation if lower levels seem inadequate to evaluate the stimuli.

As mentioned above, stimulus checks are sequential, and this sequence is roughly based upon the following steps (ignoring, again for simplicity, the last step related to normative significance). We also refer to these steps as *levels of processing*.

- Relevance detection: The stimulus is checked for novelty, innate pleasantness/unpleasantness and goal/need relevance. If it is found to be either novel, or pleasant/unpleasant or relevant to the current needs or goals of the agent, attention is directed to the stimulus (i.e., the orientation response; orienting towards the source of the stimulus) and further processing is initiated.
- Implication assessment: The stimulus is checked for its cause (what caused it), agency of the cause (who did it), its goal conduciveness (is it good for me), its discrepancy between what the agent expected and what actually happened and finally its urgency. This step needs considerable processing resources at the schematic and conceptual level, while the first step is largely operating at the sensory-motor level.
- Coping potential determination: The stimulus is checked to evaluate if the agent is able to control the stimulus or its consequences, and if the agent has enough power to actually effectuate this control (power can have many different sources like physical strength, money, friends, etc.). Finally, if coping potential is limited, the agent evaluates whether it can afford to adjust to the situation. Coping is a process that needs massive processing resources at all three levels.

Although these steps are inherently parallel and evaluated continuously, they are sequential in the sense that the later steps are only deployed to the maximum if earlier checks indicate that this is necessary. Later checks are fully activated only when earlier checks achieve “preliminary closure” (Scherer, 2001), that is, the check has to come to an intermediate stable conclusion about stimuli.

An important aspect of the SEC model is that a SEC is a continuous process that depends on, and changes the results of other SECs (including itself) and that the current state of all SECs is represented in appraisal registers (Scherer, 2001). We call this state the *appraisal state*. This state continuously synthesizes appraisal information from the SECs and is compatible with the concept of appraisal integration proposed by Smith and Kirby (2000). We do formalize the appraisal state, but we do not formalize all recursive connections between the SECs.

A second important aspect of the SEC model is that this appraisal state has a direct effect on all subsystems of the agent. For example, on information processing (the central nervous system), system regulation towards the novel situation (central nervous system, endocrine system and the autonomic nervous system) and action selection (the sensory-somatic nervous system). In the specification of the integration of both models we restrict ourselves to this

appraisal state and do not go into the details of the effects of this state on the subsystems, therefore we do not formalize the action tendencies, physiological changes and expressive behaviors that are associated with the different appraisal states.

#### 7.4.2 Smith and Kirby's ADM

We present a short overview of Smith and Kirby's ADM. In this model the appraisal state (or appraisal integration) is produced by the *appraisal detectors*. The definition of such detectors is the central feature of the ADM (Smith & Kirby, 2000). These detectors continuously integrate the appraisal results originating from three different modes of processing: stimulus perception, associative processing and reasoning. These detectors do not appraise stimuli themselves. Stimulus perception outputs appraisal information to the detectors based on the evaluation of pain, intrinsic pleasure, and other biologically important survival information. In contrast, the latter two modes are considered to be cognitive modes. Associative processing outputs appraisal information based on learned combinations of information and appraisal results. Associative processing is fast, continuous, and autonomous. It can be unconscious and is based on spreading activation paradigms. Associative processing can use any kind of information (e.g., sensations, images, sounds, and emotions). The reasoning mode outputs appraisal information based on deliberative thought processes. These processes are costly and slow, but powerful and able to re-appraise remembered situations and reflect upon the current appraisal state. Reasoning actively generates appraisal information for the appraisal detectors and corresponds best to “active posing and evaluating of appraisal questions” (Smith & Kirby, 2000). Furthermore, the more cognitive the mode, the more resources it needs. It should be clear that these modes are compatible with the levels of appraisal as described in the SEC model. Furthermore, the appraisal integration is responsible for the emotional response, which is also compatible with the appraisal registers in the SEC model.

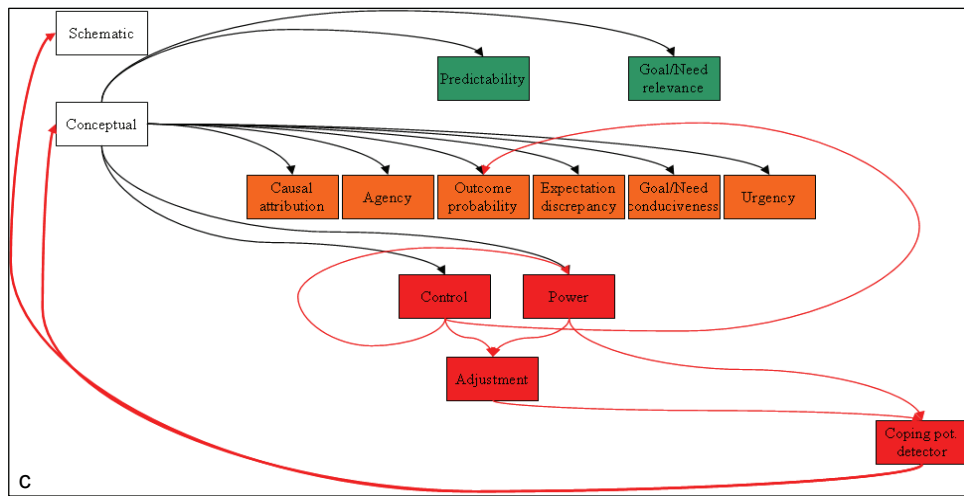
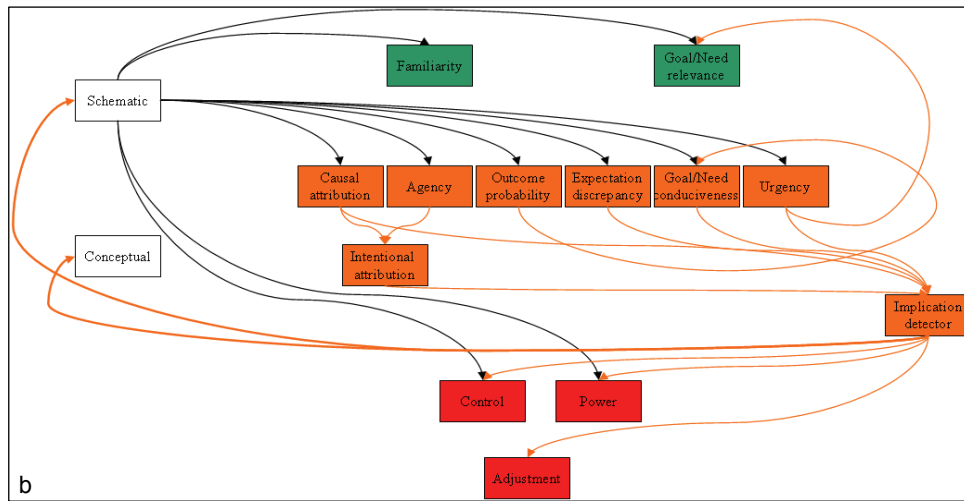
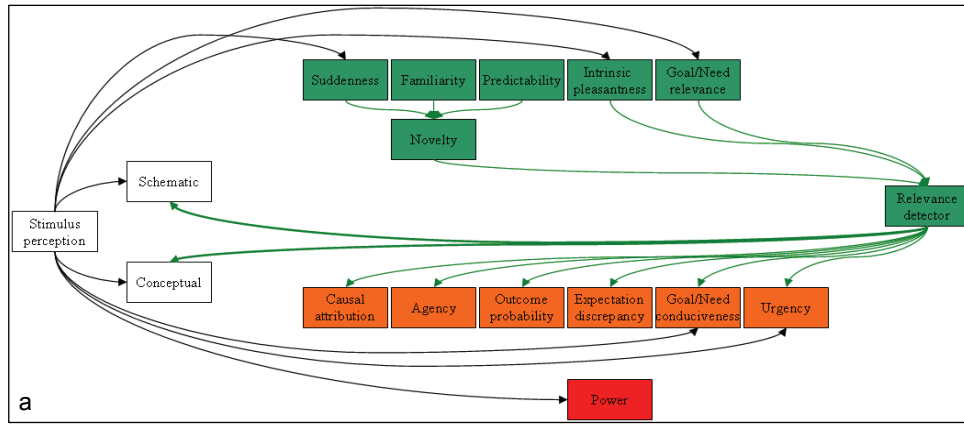
The ADM explicitly defines a feedback relation between the emotional response and the two different modes of cognitive processing. This feedback relation allows these modes to use emotional information for processing. Associative processing uses this information in learning and remembering, while reasoning uses this information to reflect upon and reappraise the situation.

### 7.4.3 Summary of Both Models

The ADM assumes three modes of information processing (stimulus perception, associative processing and reasoning). These modes generate appraisal information that is subsequently integrated into a “global” representation of the current appraisal state. This appraisal state is responsible for the emotional response. This state also feeds back to two of the three modes, namely the associative processing and reasoning modes, in order to use this emotional information for learning and reasoning respectively. The SEC model assumes three different levels of appraisal (sensory-motor, associative and conceptual) in which a large amount of different stimulus checks are present. These checks evaluate the stimuli in a specific order and depend on one another. The results of these checks are accumulated into appraisal registers, which—when the results are sufficiently stable—subsequently initiate next appraisal steps and the emotional response.

### 7.4.4 Formal Integration: the SSK Model

We now present the specification of a potential integration of the Stimulus Evaluation Check model and the Appraisal Detector Model, as an example of how our formalism can be used to integrate appraisal theories. For clarity, the specification is presented in a graphical form (Figure 7.3). To get an idea of the actual set notation, see the boxed text in Section 7.5.2 in which a simplified version of the SSK model is fully specified. This specification is used as basis for the computational model described in Section 7.5.



**Figure 7.3.** Graphical representation of the formal SKK model. See main text for explanation. Note that the boxes in the above figure denote processes. Connections between the boxes thus define process dependencies. Appraisal dimensions and emotional-response components are not represented in this figure (appraisal processes and mediating processes are, but in our formalism appraisal dimensions ( $D$ ) and processes responsible for appraising on those dimensions ( $A$ ) are not the same). Lastly, colors correspond to different appraisal steps (green: relevance; orange: implication; red: coping). Figure *a* represents all processes having incoming dependencies related to stimulus perception or outgoing dependencies related to the relevance check. Figure *b* represents the same but now for schematic reasoning or implication, while Figure *c* represents the same but now for conceptual reasoning or coping. Note that some appraisal processes receive input from all three types of processing, and as such are appraisal processes that can function on all three levels of processing (e.g., goal/need relevance).

Before describing the integrated model, some naming issues have to be resolved. When we use the term perception process, we refer to one of the three *processing modes* of the Appraisal Detector Model, to one of the three *levels of appraisal* in the SEC model and to an element  $p_i \in P$  in our formal notation. When we use the term *appraisal process*, we refer to a single *stimulus check* in the SEC model and to an element  $a_i \in A$  in our formal notation. When we use the term *mediating process* we refer to the appraisal detector/integrator in the Appraisal Detector Model, to the processes that check for *preliminary closure of the temporal appraisal result* in the SEC model and to an element  $m_i \in M$  in our formal notation.

We base our integration on two common architectural concepts of the models: (1) the separation of appraisal into three distinct levels of information processing and (2) the appraisal registers/detectors. In our integration we focus on processes (perception, appraisal and mediation) and their dependencies.

We first formalize appraisal dimensions. For clarity, we limit ourselves to the strict minimum of data to be formally specified, in our case the set of appraisal dimensions. To demonstrate the use of dependency guards with second-order conditions relating to these dimensions, we need to include in our formal description at least these appraisal dimensions. The set of appraisal dimensions is defined based on the appraisal registers described in the SEC model, excluding those related to the norm/self compatibility check:

$$D = \{novelty\_dim, intrinsic\_dim, relevance\_dim, conduciveness\_dim, \\ urgency\_dim, control\_dim, power\_dim\}$$

We continue with the perception processes. Regarding perception processes, we first define the three processing levels as perception processes, and connect these perception processes to the appraisal processes as defined by the SEC model. This is consistent with both models. The set  $P$  is represented by the white boxes in Figure 7.3 and equals:

$$P = \{\textit{stimulus perception}, \textit{schematic}, \textit{conceptual}\}$$

Second, the SEC model assumes that certain checks have input from different levels of processing. For example *Goal/Need relevance*, *Urgency* and *Power* use input from all three levels of processing. The ADM specifically assumes that appraisal information can come from different levels. Although the models do not exactly define how the appraisal processes are distributed over the three levels of processing, together they give enough guidelines to formalize the connections between perception and appraisal. These connections are shown by the black arrows in the graphical representation of the specification. These connections define *excitatory* dependencies between the perception processes and appraisal processes. This connection topology thus defines the dependencies between modes of processing / levels of appraisal on the one hand and appraisal processes on the other. Additionally two *excitatory* dependencies are defined between the perception processes: one dependency between *stimulus perception* and *schematic*, the other between *stimulus perception* and *conceptual*. This reflects the general information processing architecture of the Appraisal Detector Model, which prescribes that perceived stimuli are processed further by the associative and reasoning mode. We do not define guard conditions for these dependencies, although several exemplary guards based on the SEC model are shown in Section 7.5.2 (boxed text).

An important characteristic of both models is that appraisal processes can evaluate continuously. In our model, continuous evaluation can be initiated by the perception processes, and is independent of the previous appraisal check. This aspect is represented by the dependencies between the perception processes and the appraisal processes in the three appraisal checks. Perception processes thus influence processing of appraisal processes directly, but only according to the structural relations defined in the SEC model.

Now we formalize the appraisal processes. The colored boxes represent appraisal processes (excluding the rightmost three boxes, to which we will return shortly). The green boxes represent those appraisal processes that are part of the first step of the stimulus checking process as defined by the SEC model. The yellow boxes represent the second step and the red boxes the third step (recall that



we did not include the fourth, norm/self related step in our formal integration). The set of appraisal processes is thus defined as follows:

$$A = \{\text{elements of the set of stimulus checks in the first 3 steps of the SEC model}\} \cup \{\text{agency, suddenness, familiarity, predictability}\}$$

We have included the appraisal process *agency*, because the SEC model, when determining whether the cause of an event is due to the action of an agent, implicitly assumes the existence of this process. Also, we included *suddenness*, *familiarity* and *predictability*, the three sub checks responsible for the result of the *novelty* check. We have explicitly included these sub checks as separate appraisal processes because in the SEC model each of them operates on a different level of appraisal. Therefore, these processes need to be formally connected to different perception processes.

Connections originating from appraisal processes define *excitatory* dependencies. The topology of these connections defines the structural dependencies between appraisal processes, consistent with the SEC model. For clarity, the color of a connection represents the appraisal step to which the dependency's originating appraisal process belongs. For instance, the green connection from *suddenness* to *novelty* represents an excitatory dependency originating from an appraisal process in the first appraisal step.

We continue with formalizing mediating processes. The three rightmost colored boxes represent mediating processes. The set *M* contains the following elements:

$$M = \{\text{relevance detector, implication detector, coping potential detector}\}$$

These mediating processes are positioned between the different levels of appraisal. Mediating processes are activated by the appraisal processes of one level and activate appraisal processes of the next level, through *excitatory* dependencies. This connectivity explains their role: mediating processes detect when appraisal information is such that the next appraisal step should be activated in full glory. For example, if the *novelty* appraisal process outputs appraisal information that characterizes high novelty, the *relevance detector* will activate the appraisal processes to which its connections point.

Remember that all connections can be guarded, although for clarity we did not define most of the guards. In principle this allows connections to activate based on evaluation of second-order logic conditions. For example, we could define the following guard for the dependency between *novelty* and *relevance detector*:

$$(\exists x x \in V \wedge x = (o, d, i) \wedge i > t \wedge d = \text{novelty\_dim}),$$

with  $\text{novelty\_dim} \in D$  and  $t \in [0, 1]$  an arbitrary threshold. This guard checks the existence of a *novelty\_dim* value greater than an arbitrary threshold  $t$ . Only if this value exists, the guard will be true, and thus the connection is active. Now the *novelty* appraisal process excites the *relevance detector*.

Finally we formalize process feedback. To formally represent the influence of mediating processes on processing modes, we have defined dependencies originating from the mediating processes ending at the *schematic* and *conceptual* perception processes. The influences are represented by six thick connections between the mediating processes and the perception processes. In the ADM, the emotional response feeds back to the associative and reasoning modes. The mediating processes in our formalism generate emotional response component intensities (elements in the set  $I$ ). These component intensities formally represent the emotional response, and are available to all perception processes. Since the ADM defines this relation as data flow, perception processes are not activated through an *excitatory* dependency. We have defined a different type for these dependencies, called *information\_available*. This means that when the guard of the dependency is true, the target process is informed of the fact that new information is available.

### 7.4.3 Summary

Integration and comparison are important reasons to formalize appraisal theories (Wehrle & Scherer, 2001). Therefore, a formalism for structural models should facilitate integration and comparison. In this section we have shown how our formalism can be used to integrate theories of appraisal. We have based our integration on two common architectural concepts of the models: (1) the separation of appraisal into three distinct levels of information processing and (2) the appraisal registers/detectors. We believe the integration was greatly facilitated by the formalism's ability to describe in detail the processes, their conditional dependencies based on second-order predicates and the appraisal dimensions.

## 7.5 Using Formal Notation to Develop and Evaluate a Computational Model of Emotion

To show the power of our formalism as basis for computational models of emotion, we describe a computational emotional agent that has been based on a simplified version of the SSK model. We have emotionally instrumented an existing version of the arcade game *PacMan*. This version was downloaded from the internet (Chow, 2003). We assume that the reader is familiar with the game of *PacMan*. First, we present the specification that was used as basis for the appraisal mechanisms implemented in *PacMan*. Then we show how this specification can be used to fill in missing declarative information that is critical to the development of a computational model. Finally, we show how our formal model helped us to debug our emotional *PacMan*-agent.

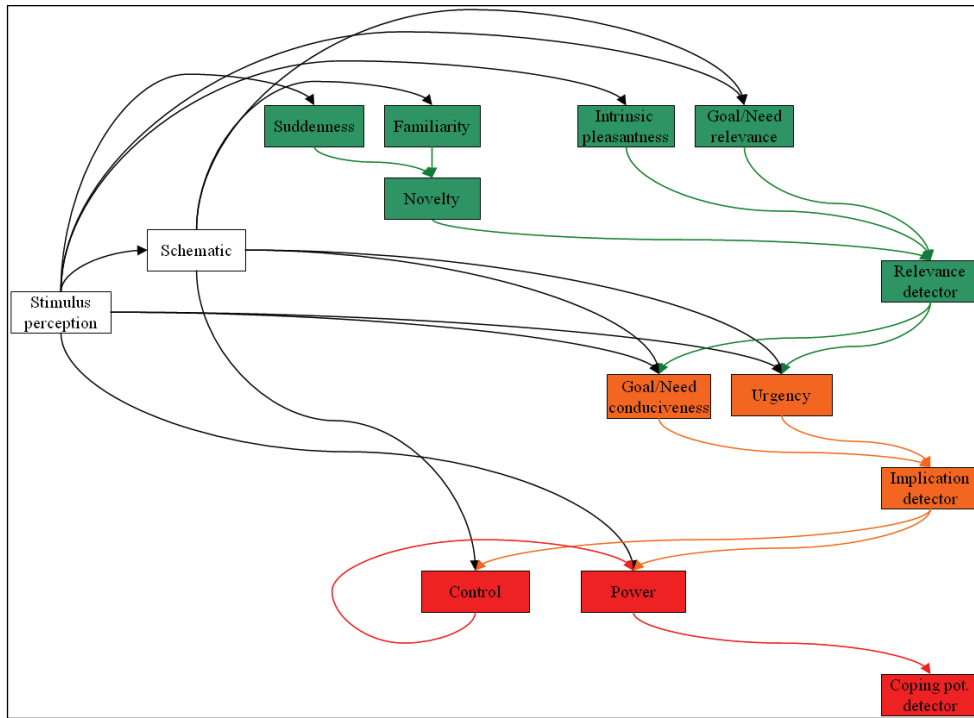
### 7.5.1 Why *PacMan*?

*PacMan*-like environments have been used in emotion research, both in the appraisal-theoretic domain (Wehrle & Scherer, 2001) as well as the virtual agent domain (Broekens & DeGroot, 2004a). Apart from being useful in the domain of emotion research (Wehrle & Scherer, 2001), *PacMan* (Figure 7.5) is also a suitable environment to test emotional instrumentation for several reasons. First, *PacMan* provides a simple environment that allows for meaningful emotional instrumentation related to different levels of appraisal. This allows us to start with appraisal processes related to sensory-motor perception only (e.g., eating dots, being eaten by ghosts) and then extend this to appraisal processes related to the schematic level (e.g., eating fruit and ghosts related to the goal of collecting points). Second, *PacMan* is an environment enabling broad emotional coverage. Many different emotions make sense. Eating ghosts, eating dots, losing a life, being chased, chasing, etc. are all different situations imbuing different emotions in humans. Third, *PacMan* is an “action-packed” environment, which allows us to test the computational model’s appraisal behavior under continuous-time constraints. This facilitates studying the process of appraisal.

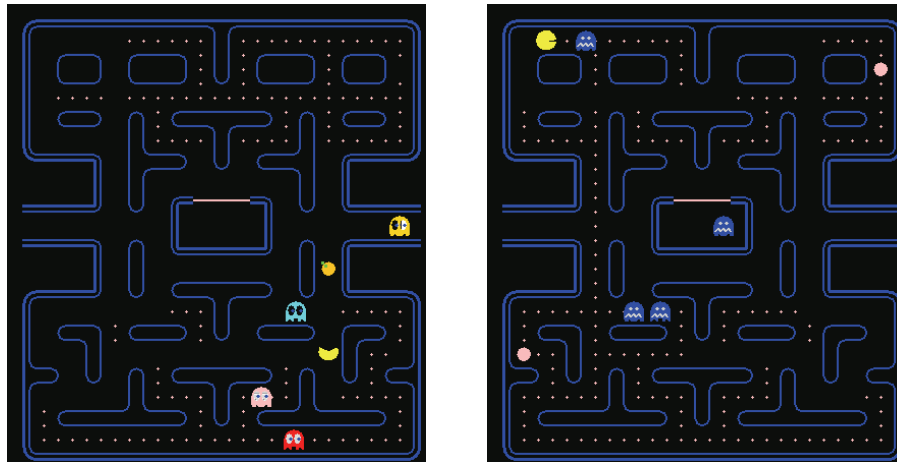
### 7.5.2 Generating a Formal Description for the Computational Model.

Before we introduce our formal description of *PacMan*’s appraisal structure we have to stress again that the point we want to make is that formal specifications of structural models are important for the development of computational models of emotion. More specific, the formal notation presented in this chapter is a powerful one. Consequently, the goal of this experiment was not to design a believable or “full-blown” emotional agent.

We have used a simplified version of the SSK model as basis for our computational emotional agent. First, we ignore the *conceptual* perception process since our *PacMan* agent is incapable of high-level cognitive processing. Second, several appraisal processes in the SSK model are ignored, because (1) these made no sense in light of the simplicity of the *PacMan* environment, or (2) because we could not design simple appraisal processes directly related to those mentioned in the formalism without providing the underlying mechanisms in more detail. Omitted processes are: *adjustment*, *expectation discrepancy*, *outcome probability check*, *predictability* and *attribution*. Third, since our *PacMan* agent is unable to use its emotions in any way, the feedback from the *mediating* processes to the *perception* processes is ignored. Note that our formal description of the SSK model enabled us to quickly evaluate what processes could or should be ignored in *PacMan*'s case. This task would have been much more difficult without such description. The resulting processes and their dependencies are depicted in Figure 7.4.



**Figure 7.4.** Graphical representation of the specification of PacMan's appraisal structure.



**Figure 7.5.** PacMan screen shots: chasing fruit (left), chasing an edible ghost (right).

The formal set notation of the simplified SSK model applied to *PacMan* is defined as follows. Perception, appraisal and mediating processes (just the processes, not the formal description of their input-output relations):

$P = \{ \text{stimulus\_perception, schematic} \}$

$A = \{ \text{suddenness, familiarity, novelty, intrinsic\_pleasantness, relevance, conduciveness, urgency, control, power} \}$

$M = \{ \text{relevance\_detector, implication\_detector, coping\_potential\_detector} \}$

Mental object types, mental objects, appraisal dimensions and emotion components:

$OT = \{ \text{belief} \}$

$PO = \{ (\text{see\_ghost, belief}), (\text{lost\_ghost, belief}), (\text{eaten\_by\_ghost, belief}), (\text{see\_edible\_ghost, belief}), (\text{lost\_edible\_ghost, belief}), (\text{eaten\_ghost, belief}), (\text{see\_power, belief}), (\text{eaten\_power, belief}), (\text{see\_dot, belief}), (\text{eaten\_dot, belief}), (\text{see\_fruit, belief}), (\text{lost\_fruit, belief}), (\text{eaten\_fruit, belief}) \}$

$D = \{ \text{novelty\_dim, intrinsic\_pleasantness\_dim, conduciveness\_dim, relevance\_dim, urgency\_dim, control\_dim, power\_dim} \}$

$E = \{ \}$

Link types, guards, data constraints and dependencies:

$LT = \{ \text{activation} \}$

$G = \{ \text{true, guard}_1, \text{guard}_2, \text{guard}_3 \}$  with:

$\text{guard}_1 = (\exists v_1, v_2, v_3 \ v_1, v_2, v_3 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge v_3 = (o, d_3, i_3) \wedge (|i_1| + |i_2| + |i_3|) / 3 > 0.15 \wedge d_1 = \text{novelty\_dim} \wedge d_2 = \text{intrinsic\_dim} \wedge d_3 = \text{relevance\_dim})$

$\text{guard}_2 = (\exists v_1, v_2 \ v_1, v_2 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge (|i_1| + |i_2|) / 2 > 0.25 \wedge d_1 = \text{conduciveness\_dim} \wedge d_2 = \text{urgency\_dim})$

$\text{guard}_3 = (\exists v_1, v_2 \ v_1, v_2 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge i_1 * i_2 > 0 \wedge d_x = \text{control\_dim} \wedge d_y = \text{power\_dim})$

$H = \{ c_1, c_2 \}$  with:

$c_1 = ((\exists x) x \in V \wedge x = (y, d, i, t) \wedge i > 0)$  if  $((\exists y) y \in O \wedge y = (c, j, t') \wedge j > 0 \wedge t = t')$ , and

$c_2 = ((\exists z) z \in I \wedge z = (e, i', t'') \wedge i' > 0)$  if  $((\exists x') x' \in V \wedge x' = (y', d', j', t''') \wedge j' > 0 \wedge t'' = t''')$

$L = \{ (\text{stimulus\_perception, suddenness, true, activation}), (\text{stimulus\_perception, intrinsic\_pleasantness, true, activation}), (\text{stimulus\_perception, relevance, true, activation}), (\text{stimulus\_perception, conduciveness, true, activation}), (\text{stimulus\_perception, urgency, true, activation}), (\text{stimulus\_perception, power, true, activation}), (\text{schematic, familiarity, true, activation}), (\text{schematic, relevance, true, activation}), (\text{schematic, conduciveness, true, activation}), (\text{schematic, urgency, true, activation}), (\text{schematic, control, true, activation}), (\text{suddenness, novelty, true, activation}), (\text{familiarity, novelty, true, activation}), (\text{novelty, relevance\_detector, true, activation}), (\text{intrinsic\_pleasantness, relevance\_detector, true, activation}), (\text{relevance, relevance\_detector, true, activation}), (\text{relevance\_detector, conduciveness, guard}_1, \text{activation}), (\text{relevance\_detector, urgency, guard}_1, \text{activation}), (\text{conduciveness, implication\_detector, true, activation}), (\text{urgency, implication\_detector, true, activation}), (\text{implication\_detector, control, guard}_2, \text{activation}), (\text{implication\_detector, power, guard}_2, \text{activation}), (\text{control, coping\_potential\_detector, guard}_3, \text{activation}), (\text{power, coping\_potential\_detector, guard}_3, \text{activation}) \}$

To construct a computational model that can execute, we have to fill in missing declarative information. We need to address several issues mentioned earlier in this chapter, issues that relate to computational aspects like process activation thresholds, process activity, and input/output constraints. Many of these questions are answered neither in the SEC model nor in the ADM. Consequently, answers are not available in the specification of the integration of both models. This is not intended as critique, but as an observation about the immediate applicability of appraisal theories as basis for computational models. This applicability is limited, as already mentioned by Gratch and Marsella (2004). Our observation lends formal support to this. We now describe how we added guards to fill in the missing details in a formal way.

First, two appraisal processes, *suddenness* and *familiarity* influence the appraisal dimension *novelty\_dim*. How does the *novelty* process integrate this information? In Scherer's SEC model (Scherer, 2001), references are made to the mechanisms that could be responsible for *suddenness* and *familiarity*, but this information is not detailed enough for a computational implementation of the integration of the results of these mechanisms. To stay consistent with the SEC mode, we assume that both *suddenness* and *familiarity* appraise mental objects in terms of the *novelty\_dim* dimension. Whenever one of these processes is active, the *novelty* check is activated and integrates these two results into one value by adding-up. Dependencies between *suddenness* and *familiarity* on the one side and *novelty* on the other are therefore without guard.

Second, what are the thresholds for the activation of the *relevance* and *implication detectors*? Or even more fundamentally, can we speak of a threshold? According to the SEC model, we can, since this model specifically mentions *preliminary closure*. However, no threshold or guideline for a threshold mechanism is given that is useful for an algorithmic approach (apart from the appraisal register values being *relatively stable*, which is about the same as *preliminary closure*).

Since we do not have a numerical guideline, we assume the following: the *relevance detector* is activated by either one of the three appraisal processes: *novelty*, *intrinsic pleasantness* and *need/goal relevance*. Every outgoing dependency from the *relevance detector* to an appraisal process of the next appraisal step has a guard equal to:

$$(\exists v_1, v_2, v_3 \text{ with } v_1, v_2, v_3 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge v_3 = (o, d_3, i_3) \wedge$$

$$(|i_1| + |i_2| + |i_3|) / 3 > 0.15 \wedge d_1 = \text{novelty\_dim} \wedge d_2 = \text{intrinsic\_dim} \wedge d_3 = \text{relevance\_dim})$$

We assume all three tuples  $v_1$ ,  $v_2$  and  $v_3$  to exist. If not, we take their corresponding activation value to be equal to 0. Thus, this guard checks the value of the cumulative activation of the appraisal dimensions that are relevant to the *relevance* check. The value must be greater than an arbitrarily chosen threshold.

The next guard is related to the *implication detector*. The Goal/Need *conduciveness* and *urgency* processes activate this *implication detector*. Every outgoing dependency from the *implication detector* to an appraisal process of the next level of appraisal has a guard equal to:

$$(\exists v_1, v_2 \text{ with } v_1, v_2 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge (|i_1| + |i_2|) / 2 > 0.25 \wedge \\ d_1 = \text{conduciveness\_dim} \wedge d_2 = \text{urgency\_dim})$$

Again, we assume that the tuples  $v_1$  and  $v_2$  exist, and if they do not, we take their corresponding activation value to be equal to 0. Thus, this guard checks the value of the cumulative activation of the appraisal dimensions that are relevant to the *implication* check.

A third missing detail is the exact relation between control and power. Also, how do these appraisal processes together influence the *coping-potential detector*? Only a descriptive guideline is given in the SEC model, stating that the evaluation of power only makes sense if the situation is controllable. Complete lack of control or complete lack of power both result in lack of coping potential. High control results in coping potential fully dependent on power. Assuming that both dimensions cannot attain negative values, this can be interpreted as a multiplication of the appraisal dimension values for *power\_dim* and *control\_dim*. *Coping potential* is activated when the product between *power\_dim* and *control\_dim* is above a certain threshold. We defined the following guard attached to the dependency between the power appraisal process and *coping-potential detector*:

$$(\exists v_1, v_2 \text{ with } v_1, v_2 \in V \wedge v_1 = (o, d_1, i_1) \wedge v_2 = (o, d_2, i_2) \wedge i_1 \times i_2 > 0 \wedge \\ d_1 = \text{control\_dim} \wedge d_2 = \text{power\_dim})$$

Again we assume that both tuples  $v_1$  and  $v_2$  exists, and if one of them (or both) do not, we take their corresponding value to be equal to 0.



Fourth, what is, in the context of *PacMan*, a sensory-motor perception process and what is a schematic perception process? According to the Appraisal Detector Model the sensory-motor mode of processing reacts to inherently pleasant and painful stimuli or facial expressions and the SEC model states that this level of appraisal relates to stimuli having to do with basic needs, available energy and direct sensory processing—like sudden movements. Both models give a clear guideline, and we think that it is feasible to use this guideline in our domain. We have done this in the following way. The sensory-motor perception process reacts to events related to the survival of the *PacMan* agent. One can think of eating dots (*PacMan* is assumed to live of dots), being eaten by a ghost and perceiving dots and ghosts (see Table 7.2). The schematic perception process reacts to events that relate to the goal of collecting points (Table 7.3).

<i>Appraisal process</i>	<i>Dimension</i>	<i>Checking criteria</i>
Suddenness	novelty_dim	Moving objects (ghosts and fruit) are evaluated equally positive and more novel than non-moving objects (pills and dots).
Intrinsic pleasantness	intrinsic_dim	Eating a dot is positive, while being eaten by a ghost is negative.
Need relevance (survival)	relevance_dim	Events related to dots and non-edible ghosts respectively have values relative to the amount of hunger <i>PacMan</i> has and the amount of lives left (hunger is simulated based on the last time <i>PacMan</i> ate a dot).
Need conduciveness	conduciveness_dim	Based on all events related to non-edible ghosts and dots.
Urgency	urgency_dim	Based on whether the event implies a moving object. Seeing a non-edible ghost is urgent.
Power	power_dim	The power-pill time left is an indication of the amount of power left.

**Table 7.2.** *PacMan* appraisal related to survival need

<i>Appraisal process</i>	<i>Dimension</i>	<i>Checking criteria</i>
Familiarity	novelty_dim	Seeing a dot is more common than seeing a ghost, and seeing a ghost is more common than a power-pill which again is more common than fruit.
Goal relevance (points)	relevance_dim	All events related to fruit and eating ghosts are equally relevant.
Goal conduciveness (points)	conduciveness_dim	Seeing an edible ghost, eating a ghost, seeing and eating fruit are positive, while losing an edible ghost and losing a fruit are negative.
Urgency	urgency_dim	Based on whether the event implies a moving object. Seeing an edible ghost and a fruit both are equally urgent.
Control	control_dim	Based on whether the event allows to be controlled. All moving objects allow control to a certain degree, but fruit and edible ghosts allow for more control than non-edible ghosts. Seeing a power-pill also implies control.
Power	power_dim	Power is completely determined based on the power-pill time that is left.

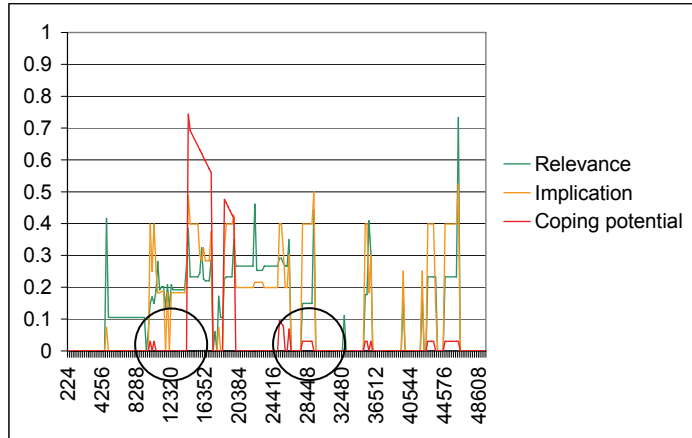
**Table 7.3.** *PacMan* appraisal related to the goal of gathering points

### 7.5.3 Verification of the Computational Model

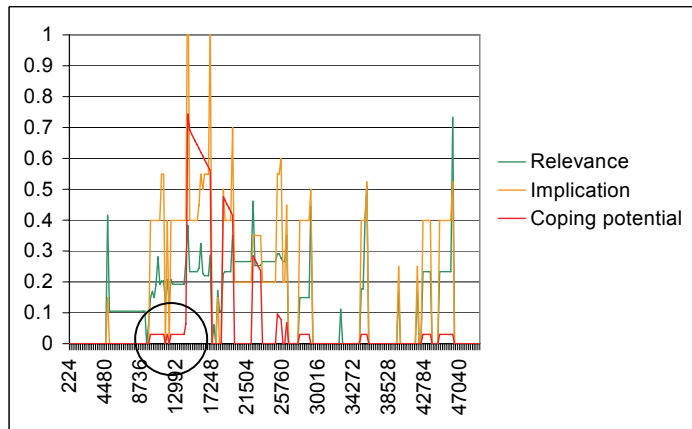
We have instrumented *PacMan* by building a simple system that generates mental objects based on the current game situation. The decision support system is based on the SSK model and has two processes, the *sensory-motor* perception process and the *schematic* perception process. Mental objects are appraised based on the appraisal processes and their relations as described in the SSK specification. These appraisal processes produce appraisal dimension values, as specified in Tables 7.2 and 7.3. These values are continuously integrated and the result is maintained in an appraisal state that is modeled as a vector with cardinality equal to the number of different appraisal dimensions (7 in our case, see boxed text in Section 7.5.2). This integration simply consists of adding appraisal values that belong to the same appraisal dimension and storing the result in the appraisal state.

The experiment itself consists of a human player controlling the instrumented *PacMan* agent who plays the first level of the *PacMan* game (by eating all dots), loses a life two times during the game, and eats several ghosts. When we ran the experiment, the result was contradictory. Although certain situations obviously should have a strong implication to the *PacMan* agent, the stimulus checks of the coping appraisal step were not activated, but should have according to the formal description. This lack of activation can be seen in Figure 7.6a, for  $9000 \leq t \leq 13000$ . In these situations *PacMan* was seeing a ghost and seeing and eating dots. However, the *implication* in this situation is below the arbitrarily defined

threshold of 0.25, while other clearly less important situations are above this threshold (e.g., around  $t=27000$  where *PacMan* only sees a ghost).



**Figure 7.6a.** PacMan using bi-polar variables. Time in milliseconds is on the x-axis. Appraisal dimension activation is on the y-axis. *Coping potential* is not activated around  $t=10000$ . The *implication detector* stays below its threshold.



**Figure 7.6b.** PacMan without bi-polar variables. *Coping potential* is activated around  $t=10000$ , as a result of higher activation of the *implication detector*.

We can explain these contradictory results by examining the formal specification. The appraisal process *conduciveness* can produce both positive and negative appraisal values for the appraisal dimension *conduciveness\_dim*. When these values are integrated by the *implication* mediating process, they cancel each other instead of together contributing to a high implication situation. Subsequently the guard of the *implication* mediating process is not true, so the next appraisal step (coping) is not activated, resulting in the contradictory result.

The underlying reason for this is that the above mentioned appraisal dimension is bi-polar (i.e., can have negative and positive values) and thus

switches meaning when it switches sign. Consequently there is only a small difference between, for example, a situation in which highly conducive and non-conductive events happen and a situation in which nothing happens at all. In other words, this dimension cannot represent “mixed-emotions”. Because of the formal structural specification of the SSK model, we were able to exactly identify this issue. After the introduction of an extra appraisal dimension, an extra appraisal process that checks stimuli related to *non-conduciveness*, and a link between that process and *implication*, the new results are as expected. *Coping potential* is activated but low since *PacMan* has not eaten a power-pill recently (Figure 7.6b, between  $9000 \leq t \leq 13000$ ).

#### 7.5.4 Summary

Our formalism helped to develop a computational model based on the SSK model. It facilitated (1) filling in of computational details, and (2) making computational assumptions explicit. Further, the formal description helped us to verify and validate our computational model with respect to the SSK model. We could identify what was in our case a problem using bi-polar appraisal dimensions. Note that we do not claim anything about bi-polar appraisal dimensions per se. We claim that our formalism is useful for the specification and verification of a computational model.

### 7.6 Discussion

We first discuss several formalization issues. Then we discuss related and future work.

#### 7.6.1 Some Drawbacks of Formalization.

Two warning remarks regarding formalization have to be made. First, the focus on strict definitions can be a disadvantage of formalization when used as a tool for psychological theory refinement. Formal modeling forces a theorist to commit to certain definitions for the concepts in a theory. In and of itself such commitment can be an advantage because it helps to refine and clarify theories (Mallery, 1988). However, such commitment can also be a disadvantage when unclear bounds of the concept to be formalized result in either a too strictly formalized concept—producing a formal representation that does not cover all of the concept—, or a too loosely formalized concept—producing a formal representation that is not better than the non-formal representation. It could be argued that this is not a disadvantage of formalization, but a lack of specificity of the theory. The theory lacks clear definitions. However, appraisal theories—like

many theories of psychological processes—generally include concepts with such open bounds for good reasons.

A second, more important, disadvantage is that formal specifications risk living their own lives. This is all right if the probability is high that a formal specification covers everything the theory describes. As discussed above, exactly *this* is far from certain. However, as formal notations have many benefits (clarity, preciseness, etc.) the formal description of a cognitive appraisal theory might (by some) be interpreted as a substitute for the actual theory. This could result in overly strict interpretations of that theory, eventually leading to wrongly rejecting a phenomenon as consistent with the theory, based on results from an experiment with a computational model that is based on a formal specification. Rejecting a phenomenon based on a formal description of a psychological theory should thus always be done with care. The inverse, the acceptance of a phenomenon as supporting a theory, is less problematic since the formal specification of the theory generally is stricter than the theory itself.

### 7.6.2 Related Work.

We briefly discuss four approaches to the formalization of emotion theory. The choice for these four examples is not arbitrary; they each represent a different way in which formalization can be used in this context.

First, Gmytrasiewicz and Lisetti (2002) have defined a formalism to describe how emotions can influence agent decision making. Their formalism defines emotions as different modes of decision making. Their formalism allows the definition of personalities of others, where a personality can be seen as the potential transitions between emotional states. This approach is different and in a way complementary to ours. While their approach takes the emotion as a given and formalizes the influence this emotion has on decision making, our approach formalizes the structure of appraisal in order to, for example, describe the interactions between perception, appraisal and emotion mediating processes that generate the emotion in the first place.

Second, Meyer (2004) proposes a formalism based on modal logic to formally describe how specific emotions relate to the belief, desire and intention structure of an agent. This approach differs from ours in the sense that it tries to formalize an emotion in terms of specific sets of beliefs, desires and intentions, while our approach tries to formalize the appraisal theory on which the computational model is based by describing the processes and their structural relations.

Third, the GATE environment is a black-box modeling environment aimed at theory comparison (Wehrle & Scherer, 2001). This tool allows researchers to specify the theoretical relation between appraisal dimension intensities and emotional-response components—using mathematical formulas and parameters—and quickly compare the results of experiments with the theoretical predictions. A large database is attached to the tool, in which experimental results are stored. The database can be filled automatically with the results of questionnaires that are filled in by subjects. Data from this database can be used to compare experimental data with theoretical predictions derived from various theories. GATE contains a large set of analysis functionality to facilitate this comparison. The main differences between GATE and our approach are our theory independent, set-based formalism and our focus on the specification and verification of computational models. Our formalism allows the definition of the declarative semantics of the different processes, their inputs, outputs and interactions. If time is introduced (see future work) in our formalism it enables specification of the relation between the sub-processes involved in appraisal and specification of evolution of the structure of appraisal during development of an agent. Since we use a set-based notation, a formal specification developed with it can be systematically and automatically evaluated for consistency with a computational model or appraisal theory.

Fourth, Reisenzein (2000) proposes a meta-level formal representations for the emotion theory of Wundt. His approach is very similar to ours, in that it attempts to formalize the emotion theory at a structural level using a set-theoretic notation. Important differences are that his approach is more systematically based upon the structuralist approach (Westmeyer, 1989), and that our formal notation has explicitly been developed to also facilitate development of computational models. However, a closer comparison of both approaches is needed in the future. This is specifically interesting as the structuralist approach towards formalization is by no means restricted to the formalization of cognitive theories. This would indicate that our approach could be extended to less cognitively-oriented theories of emotion.

### 7.6.3 Future Work.

Our current version of the formal notation describes the static structure of appraisal. Future work should include time. Time is needed in order to model the evolution of a structural model. For example, we might want to formalize the relation between different developmental stages from child to parent (Lewis, 2001), or formalize the evolution of an appraisal over a shorter time period.

Further, to formalize the difference between conscious and unconscious influences (Zajonc, 2000), we need to separate the mental objects, our set  $O$ , in subsets of objects. Every subset now contains objects with different activation strength. This strength represents whether an object is conscious or not.

Also, future work includes the addition of long term memory to our formalism. It is difficult to formalize reappraisal (Levine, Prohaska, Burgess, Rice & Laulhere, 2001) or coping (Lazarus, 2001), without the LTM construct.

Finally, a comparison between the structuralist approach towards theory formalization and our approach is planned.

## 7.7 Conclusion

Integration of appraisal theories is important for the advancement of appraisal theory (Wehrle and Scherer, 2001). We have proposed a formal notation for the declarative semantics of the structure of appraisal, and argued for the need to have such a formalism. We have shown that this formalism facilitates integration between appraisal theories. We have illustrated this by integrating (in a simplified way) two appraisal theories; the Stimulus Evaluation Check model by Scherer (2001), and the Appraisal Detector Model by Smith and Kirby (2000) into one model, the “SSK model” (Section 7.4). The process of integration was greatly facilitated by the ability provided by the formalism to specify in detail the perception, appraisal and mediating processes, their conditional dependencies based on second-order logic and the appraisal dimensions.

We have shown that our formalism is a first step to narrow the gap between structural models of appraisal and computational models. To this end we have used our formalism as intermediate specification of structure and completed the translation process from appraisal theory to computational model by developing a computational model of emotion based on the “SSK model”. We have shown that our formalism helped development in the following way (Section 7.5): filling in of computational details, and making computational assumptions explicit was greatly facilitated by the formal description of the “SSK model”. Moreover, it helped us to verify and validate our computational model with respect to the “SSK model”.

To summarize, our formalism for the structure of appraisal can be used to further advance cognitive appraisal theory as well as to facilitate development and evaluation of computational models of emotion based on cognitive appraisal theory.