Running head: Formal Models of Appraisal.

Formal Models of Appraisal: Theory, Specification, and Computational Model

Joost Broekens[1],

Doug DeGroot,

Walter A. Kosters,

Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands.

Correspondence to:

Joost Broekens

Niels Bohrweg 1

2333CA, Leiden

The Netherlands

Phone: +31 (0)6-24657367

Fax: +31(0)71-5276985

Email: joost.broekens@gmail.com

---

[1] Currently at the Telematica Institute, Enschede, The Netherlands.

Abstract

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 paper 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.

Keywords: Affective Computing, Cognitive Appraisal Theory, Computational Modeling, Emotion, Formalism, Theory Analysis.

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). 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 and 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 article.

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 article 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 article we propose a formalism to describe the structure of appraisal (Section 3) and we elaborate on two types of uses for this formalism: (1) we use it to integrate two different appraisal theories (Section 4), and (2) we use it to analyze a computational model of emotion we developed (Section 5). Before continuing the main line of this article, we first give a cognitive definition of emotion, some more explanation on the development and use of artificial emotional agents and a more detailed description of the problem we address.

Emotion. In humans, the term emotion refers to a set of naturally occurring phenomena including facial expression, physiological changes, motivation, actions such as fight or flight behavior, a tendency to act, feelings and cognitive appraisal (see, e.g., Scherer, 2001; Roseman and Smith, 2001). An emotional state is the combined activation of instances of a subset of these phenomena, e.g., *angry* involves a tendency to fight, a typical facial expression, a typical negative feeling, increased heartbeat, etc., while *elated* involves positive feelings, increased heartbeat, smiling, etc. Such a state is often a short-lived, intense reaction to a certain situation. Additionally, emotions are important in non-verbal communication, and emotions influence cognition in many ways (how we process information, our attention, and our biases towards information).

Many different definitions of emotion exist (Picard, 1997; Hudlicka, 2004), as well as many different theoretical views. However, in this article we focus on the cognitive view. In cognitive psychology, emotion is strongly related to goals, needs, desires and concerns of an individual (we refer to these four terms as *goals*). Emotion is elicited by the evaluation of an event in relation to the accomplishment of the agent's goals. For example, if someone wants to earn a Master’s degree and finally passes for the last exam after having tried to pass several times, a cognitive theory of emotion would predict happiness and relief immediately after hearing the positive result: the individual’s goal (the degree) has been met by passing the exam (an event), as the exam was a difficult one for which the individual has failed several times, there’s also the feeling of relief as a probable but undesirable outcome has not been realized (e.g., Ortony, Clore & Collins, 1988). An emotion can therefore be seen as a “heuristic” that mediates between events and goals (Oatley, 1999). On the one hand it is a (non-cognitive) result of cognitive contemplation, on the other the emotion itself helps us detect how we are doing in relation to our goals.

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), virtual agents in VR training environments (Henninger et al., 2002), and agents targeted at decision-making and planning (Coddington and Luck, 2003). 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. Subjects were interrupted during a problem-solving task with a computer by artificially induced mouse delays. After this interruption, the emotional intervention was given in the form of positive or negative affective speech. Positive words resulted in smiling as well as better problem solving performance.

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 and 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, but cognitive appraisal is an important component of emotion. Although previously most appraisal theories assumed that appraisal was a necessary and sufficient condition for emotion (Roseman and Smith, 2001), currently it is seen as an important component of emotion.

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 and 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 nearing 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 and DeGroot, 2006), just like we need to ask a person about why he/she does something instead of *only* looking at neuroimaging data.

Advancement of 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 same basics assumptions, many different appraisal theories exist (Reisenzein, 2001; Frijda and 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 and Scherer, 2001). Second, computational modeling of emotion is a powerful way of analyzing appraisal theories in a formal way (Wehrle and Scherer, 2001). Formal descriptions facilitate the evaluation of computational models, thereby contributing to the analysis of appraisal theories.

Aim and scope of this article. The main contribution of this article 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.

1. 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 and 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).
2. 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 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 article 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 3. While the introduction of Section 3 is essential for understanding the rest of the article, the parts that detail the formalism are recommended to the mathematically oriented reader. Less mathematically oriented readers will find Section 4—showing how the formalism can be used as a tool to facilitate the integration of appraisal theories— as well as Section 5—demonstrating how a formal description of a structural model of appraisal can be used as basis for a computational model—more interesting. Section 6 discusses issues around formalization, and related approaches. Appendix *A* presents in full the formal description of the appraisal structure used in our computational model.

2 Appraisal Theory: Structure, Process and Computation

A common classification of CATs is based on a structural versus a process-based description (Roseman and 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 evaluate the perceived environment in terms of values on a set of measures called *appraisal dimensions*. An appraisal dimension influences emotion and can be considered as a variable—e.g., agency or intrinsic pleasantness—, used to express the result of the appraisal of a perceived object—e.g., a friend. The variable itself represents a generic type of evaluation, specific to the appraisal theory at hand, while the value attributed to it due to evaluation of a situation is subjective. This process of evaluation is called *appraisal.*
- the processes that relate appraisal 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 (however, see the general discussion). In this paper 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 paper, 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 1 about here)

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 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):

- the nature and level (van Reekum, 2000) of processes; deliberative, automatic, innate?
- 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/produce?
- 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?
- the number of different appraisal dimensions, their activation range and the responsible processes.

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 the formalism in detail as well its use. Our formalism is set-based and uses three main constructs: *processes*, *data*, and *dependencies*. Following generic cognitive appraisal theoretic assumptions, three types of processes exist; *perception* processes (`P`), *appraisal* processes (`A`) and *mediating* processes (`M`) (Figure 2). All processes are formalized in the same way, i.e., as sets of functions. Such a function assigns elements from an "output" set to elements from an "input" set. The input and output are powersets that are process-type specific, for example, the input for any perception process is a combination of what happens in the world, the current appraisals and the current emotion. The output of a perception process is a mental representation of the input. The notation used for these three types of processes and the accompanying terminology are borrowed from Reisenzein (2001).

Several types of data exist, all defined as sets. The *external world,* `PW`, a set defining the *observables*, is the set of all potential events and objects that can respectively occur and reside in the environment. The set `W` is a subset of `PW` and contains the objects and events that are currently available to the agent for perception, i.e., its range of perception. Perception processes map information from the external world, as defined in the set `PW`, to *mental objects*—representations of the external world suitable for appraisal. These *internal representations* of things happening in the world (`PO`) are defined as a set of *potential percepts*, i.e., those things that can potentially be perceived. A subset of the set `PO` is the set `O`, containing all currently perceived things. This fourth set, `O`, is our formalisms representation of "working memory". We assume that, in order for an appraisal to happen, the object of appraisal has to be represented internally, which is why we label `O` as working memory (however, working memory is much more complex, and we do not claim to model working memory per se). Appraisal processes evaluate the content of working memory (the set `O`) produced by the perception processes and assign a combination of appraisal-dimension values to objects in working memory. The *appraisal dimensions* that can be used in the process of appraising a situation are defined as the set `D`. This set defines which appraisal dimensions exist according to the theory being formalized. As mentioned, appraisal processes evaluate in terms of appraisal dimension values. The data types used here are potential *appraisal-dimension values*, the set `PV`, and actual appraisal-dimension values, the set `V`. Potential appraisal-dimension values are needed to, e.g., define the range of values that can be attributed to appraisal dimensions. Mediating processes relate appraisals to emotions. In our formalism, mediating processes relate appraisal dimension values, as defined in the set `PV`, to emotion-component intensities. The data types used for this are *emotion components*, the set `E`, potential emotion-component intensities, the set `PI`, and actual emotion-component intensities, the set `I`. The set of emotion components, `E`, defines, according to the appraisal theory being formalized, the elements of emotion, e.g., facial expression of anger, arousal, the feeling of joy, etc. The sets `PI` and `I` define, analogous to `PV` and `V` for the appraisal dimension values, the range of potential emotion-component intensities and actual emotion-component intensities.

Note that, the set `D` (appraisal dimensions) and the set `E` (emotion components) apart, our formalism actually only defines four types of data: events, mental objects, appraisals, and emotion components (Figure 2) However, every such type needs two sets, one to define potential elements (so that processes can be defined as mappings from an input domain to an output range), the other to define actual elements (so that constraints can be defined whether or not certain things are actually observed, appraised etc.). This aspect will be clarified in Sections 3.1-3.3.

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 highly simplified “baby” can be defined. The baby can be exposed to a barking dog or the mother of the baby, resulting in different emotions.

(Figure 2 about here)

3.1 World, perception processes and objects of appraisal

Definition 1.1: $PW=\{w_1,\dots,w_n\}$ is the set containing potential observable objects and events in the environment of the agent[1]. The event w_1 can represent, e.g., “wind”, while w_2 represents “rain”. The set PW is the set of *observables*, i.e., the things the agent can potentially observe, if it is looking at them, being close, etc. In normal wordings this means that PW defines the world of the agent.

Definition 1.2: $W\subseteq PW$ is the set containing the objects and event that are currently available to the agent for perception. As such, W defines the external state of the agent. This is not *necessarily* the set of objects and events that are actually perceived, but the set the agent *could* perceive if there is a suitable perception process available.

Definition 1.3: $PO=\{po_1,\dots,po_n\}$ is the set containing all potential percepts. The set PO is thus defined analogous to the set PW, but contains potential “mental representations”, not observables. Obviously, there does not need to be a one-to-one mapping between the observable world PW and potential “mental representations” PO. For example, both events “wind” and “rain” might be mentally represented as “autumn”.

Definition 1.4: $O\subseteq PO$ is the current content of working memory, that is, the currently perceived objects. The set PO is thus the superset containing all potential mental objects. The relation between O and PO is the same as the relation between W and PW: O and W are subsets containing currently available elements. Depending on the appraisal theory that is formalized, additional subsets of PO can be defined. For example, a theory that defines beliefs, desires and intentions can define subsets $BELIEF\subseteq PO$, $DESIRE\subseteq PO$ and $INTENTION\subseteq PO$ to discriminate between objects that are beliefs, desires and intentions respectively. Defining such subsets introduces the need to define copies of

---

[1] Note that we use n as a finite but arbitrary number to denote multiple elements in a set. Two sets both defined with n elements do not *necessarily* have the same number of elements. When they do, another subscript is used, e.g., m. Also, $Pot(S)$ is used to denote the powerset of set S, i.e., the set of all subsets of the (usually finite) set S.

potential mental objects in $\mathtt{PO}$, *if* “the same” object can sometimes be a belief and sometimes a desire. For example, $\mathtt{PO}$ needs to contain “rain” (belief) and “want rain” (desire).

Definition 1.5: If we assume $\mathtt{PV}$ as the set of potential appraisal dimension values (see Definition 2.2) and $\mathtt{PI}$ as the set of potential emotional-response-component intensities (see Definition 3.2) then we define $\mathtt{P}=\{\mathtt{p}_1,\ldots,\mathtt{p}_n\}$ as the set of all perception processes available to the agent, with for each $\mathtt{p}\in\mathtt{P}$, $\mathtt{p}:\mathtt{Pot}(\mathtt{PW}\cup\mathtt{PV}\cup\mathtt{PI})\rightarrow\mathtt{Pot}(\mathtt{PO})$. In words, a perception process $\mathtt{p}\in\mathtt{P}$ 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 zero or more mental objects that are then part of working memory[2]. We assume the following constraint: $\forall \mathtt{o}\in\mathtt{O}\ \ \exists \mathtt{p}\in\mathtt{P}\ \ \exists \mathtt{x}\in\mathtt{Pot}(\mathtt{W}\cup\mathtt{V}\cup\mathtt{I})$ such that $\mathtt{o}\in\mathtt{p}(\mathtt{x})$, and, if $\exists \mathtt{p}\in\mathtt{P}\ \ \exists \mathtt{x}\in\mathtt{Pot}(\mathtt{W}\cup\mathtt{V}\cup\mathtt{I})$ such that $\mathtt{o}\in\mathtt{p}(\mathtt{x})$ then $\mathtt{o}\in\mathtt{O}$. In other words, $\mathtt{O}$ is the union of all $\mathtt{p}(\mathtt{x})$, where $\mathtt{p}$ varies over $\mathtt{P}$ and $\mathtt{x}$ varies over $\mathtt{Pot}(\mathtt{W}\cup\mathtt{V}\cup\mathtt{I})$. Thus, if an object is in working memory then there must be a responsible perception process and input mapping to that object, and vice versa. This constraint (and equivalent ones in Definitions 2.4 and 3.4) is needed to exclude “hallucination” (e.g., spontaneous generation of objects in $\mathtt{O}$) and “willful blindness” (e.g., non-existence of an object in $\mathtt{O}$, while the process does map available input to that object).

Note that the input of a perception process is defined over $\mathtt{Pot}(\mathtt{PW}\cup\mathtt{PV}\cup\mathtt{PI})$, while the constraint is defined over $\mathtt{Pot}(\mathtt{W}\cup\mathtt{V}\cup\mathtt{I})$. This is exactly what is needed: perception processes must be defined for *potential* events (and appraisals and emotions), while the constraint enforces a link between *actual* events and a representing working memory object.

Also note that the formalism is timeless, i.e., processes do not produce anything and there is no evolution of an internal state as processing continues; processes define static relations between (elements in) sets. Adding time, to, e.g., allow appraisals to develop over time is at this point considered future work.

---

[2] 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.

We define input and output as powersets for the following reason. To specify a certain perception, a mapping of any combination of elements in $PW \cup PV \cup PI$ to any combination of mental representations of elements in PO might be needed. Further, in our case, using a generating set instead of a powerset construction is less appropriate, as this would generate a set containing complex (combined) elements instead of subsets containing combinations of elements. Combined elements (but not subsets) are problematic for the definition of the behavior of the appraisal processes, as for each combined element a specific mapping must be specified, while for the subset, behavior can be defined for the elements in that subset. In the latter case, appraisal processes can still map objects to appraisal dimension values, even though the object appears in working memory (O) in combination with other objects. The same holds for appraisal processes and mediating processes, as defined later.

Note that our definition of perception is broader than is usual in psychology. Perception in our case is defined as the set of processes responsible for the appearance of mental representations in working memory. As such, a perception process can also be a process that remembers something, or a process that deduces a new fact from previously perceived facts, i.e., we take perception as shorthand for *perception or inference*. This enables, e.g., also remembered objects to be appraised. If a specific theory needs different types of perception processes, subsets of P can be defined, e.g., $INDUCTIVE\_REASONING \subseteq P$.

Also note that perception processes perceive the agent's current appraisal dimension values and current emotion components. These two kinds of information are mapped to mental objects. In our formalism, only perception can “fill” working memory, therefore the emotion-component intensities, I, and appraisal information, V, must be perceived (mapped to a mental representation) 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 mental representations. 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 formal representation of “being aware of your current emotional state” in contrast to “being influenced by your current emotional state" and as such facilitates explicit specification of coping, re-appraisal and strategic use of emotions. This is important for the completeness of our formalism.

In our baby example the world contains two events: the *mom of the baby entering the room* and a *dog that barks*, represented by two distinct noise levels m and d, i.e., the set $PW=\{m,d\}$. Suppose that the baby's direct surroundings only contain the dog, so $W=\{d\}$. The set of potential mental objects (i.e., the thing the baby knows of) is $PO=\{po_m,po_d\}$. The baby can perceive these events with her only perception function called hear, p_h, that perceives noise levels m and d. $P=\{p_h\}$, such that $m\in x$ for some input x if and only if $po_m\in p_h(x)$, and $d\in x$ if and only if $po_d\in p_h(x)$. Thus, p_h maps m and d (*mom* and the *dog*) to the corresponding potential mental objects po_m and po_d respectively, both elements in the set PO. Following the constraint defined in Definition 1.5, and the definition of our perception function p_h, the baby must now have only a representation of the dog, i.e., po_d, in her working memory: $O=\{po_d\}$. The baby perceives the dog.

3.2 Appraisal processes, appraisal dimensions and values

Definition 2.1 $D=\{d_1,\dots,d_n\}$ is the set of appraisal dimensions. This set contains elements that represent appraisal dimensions depending on the theory being formalized, for example, suddenness and intrinsic pleasantness.

Definition 2.2: $PV=\{v_1,\dots,v_n\}$ is the set of potential appraisal dimension values with for each $v\in PV: v=(o,d,r)$ with $o\in O$, $d\in D$ and $r\in[-1,1]$. So, $v\in PV$ is a tuple of a one-dimensional appraisal result attributed to one mental object, in other words, $v\in PV$ is the result of appraising an object in terms of one appraisal dimension. As such, we adopt the standard appraisal theoretic assumption that appraisals are object-directed, explaining the presence of an object as part of the tuple defining elements of PV.

Definition 2.3: $V\subseteq PV$ is the set of *current* appraisal dimension values, again defined in the same nature as W and O. However, as V contains tuples that include an intensity, we define the following extra constraint: if $v=(o,d,r)\in V$ and $w=(o',d',r')\in V$ with $v\neq w$ then $o\neq o'$ or $d\neq d'$. This enforces that there can only be one appraisal per object-dimension active at the same time. For explanation see Section 3.3.

Definition 2.4: The set $A=\{a_1, \ldots, a_n\}$ is defined as the set of appraisal processes, with for each $a \in A$ $a: Pot(PO \cup PI) \rightarrow Pot(PV)$ such that $\forall v \in V\ \exists a \in A\ \exists x \in Pot(O \cup I)$ with $v \in a(x)$, and if $\exists a \in A$, $\exists x \in Pot(O \cup I)$ with $v \in a(x)$ then $v \in V$. Again in words, $a \in A$ 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[3]. If an appraisal exists, a responsible process must exist with corresponding input, and vice versa. Appraisal can be biased by the current emotion, explaining I in 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.

Appraisal processes are not defined in the same way as perception processes. Perception processes map to mental objects, without using a tuple that includes the input argument as output, as is the case in appraisal processes. This difference stems from the following. If an appraisal process re-appraises a situation at time $t+1$, it needs to be able to use a mental representation o_{t+1} that is the result of the perception of the to-be-re-appraised object, and as such a perception process needs o_t as part of $v \in V$ as its argument (with $o_t \neq o_{t+1}$). If this would not be the case, the identity of the appraised object would be lost in the second appraisal round. However, there is no need to have events be a part of the to-be-appraised structure, as the mental object is the only thing that can be appraised: humans neither

---

[3] 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. One might perceive several things that influence appraisal of only one of those things, e.g., the baby appraises the dog as pleasant when the mother is near, but appraises the dog as unpleasant when the mother is gone. In this example, the mother is not appraised, but is part of the input for the appraisal process appraising the dog. Second, some theories of appraisal handle cases in which misattribution of an appraisal to an object occurs. In this case the dimension-value tuple would not at all be attached to the object being appraised. For example, the baby sees the mother *and* the dog, but appraises the mother *as* the dog.

appraise the thing "out there" but the representation of that thing, nor appraise a memory but an active representation of that memory.

To continue our baby example, the baby has two appraisal processes, intrinsic pleasantness and suddenness. Both assign tuples of values $[-1,1]$ and appraisal dimensions to mental objects. There are two appraisal dimensions with the same name as the appraisal processes. Thus $A=\{a_p,a_s\}$ and $D=\{pleas,sudd\}$. The set PV is filled with all possible combinations for (o,d,r) tuples. The dog produces noise, and we assume the baby appraises this noise as arousing and unpleasant. So, $a_p(\{po_d\})=\{(po_d,pleas,-1)\}$ and $a_s(\{po_d\})=\{(po_d,sudd,1)\}$. For all other inputs x, $a_p(x)=\emptyset$ and $a_s(x)=\emptyset$. The set V equals $\{(po_d,pleas,-1),(po_d,sudd,1)\}$ as $O=\{po_d\}$.

3.3 Formalizing the mediating processes

Definition 3.1: $E=\{e_1,\ldots,e_n\}$ is the set of possible components of the emotional response, like certain subjective feelings, facial expressions, physiological reactions and action tendencies.

Definition 3.2: $PI=\{i_1,\ldots,i_n\}$ is the set of emotional-response-component intensities with for each $i\in PI$ with $i=(e,r)$, $r\in[-1,1]$ and $e\in E$. In words, $i\in PI$ is the intensity of one specific emotional-response-component, e.g., a heart-rate of 0.5 (on some scale). Subsets of PI are called *complexes*, and constitute what can be called an emotion (at least all parts of an emotion apart from the appraisal). How a complex looks like is again defined based on the theory at hand. For example, $JOY\subseteq PI$, with joy containing elements such as $(happyfeeling,\ r),(smile,\ r),(aroused,\ r)$ with $r>0.5$. Note that the set PI is object-less. Most appraisal theories assume that appraisal dimension values instead of emotional-response-component intensities are attributed to objects. As a result, in our formalism the appraisal is object-directed, but not the (other) emotional components.

Definition 3.3: $I\subseteq PI$ is the set of *current* emotion-component intensities, again defined in the same nature as W, O, and V. As I contains tuples that include an intensity, we define the following extra constraint: if $i=(e,r)\in I$ and $j=(e',r')\in I$ with $i\neq j$ then $e\neq e'$.

Definition 3.4: $M=\{m_1,\ldots,m_n\}$ is the set of mediating processes with for each $m\in M$ $m:Pot(PV)\rightarrow Pot(PI)$ such that $\forall i\in I\ \exists p\in P\ \exists x\in Pot(V)$ with $i\in p(x)$, and if $\exists p\in P$, $\exists x\in Pot(V)$ with $i\in p(x)$ then $i\in I$. In words, a mediating process $m\in M$ maps appraisal dimension values to emotional-response-component intensities such that every emotion-component intensity has a corresponding input and vice versa.

Note that a process $m\in M$, in contrast to a process $a\in A$, does not map its input to an object. This also follows a common appraisal conception that appraisals are directed at objects, but emotions are objectless. Therefore, the object is no longer needed in the result of the mediating process. Should a specific theory need to specify perception or appraisal influence by emotion, this is possible in our formalism, as PI (potential emotion-component intensities) is part of the input for P and A.

The output of appraisal processes and mediating processes are sets of tuples that include intensity. This introduces two problems. If, e.g., two appraisal processes would produce the same (o, d, r) tuple then only one can be represented in the set V (per the definition of sets). However, this would mean that the total intensity of the appraisal dimension cannot be represented by our formalism (two elements that are the same cannot both be represented in the same set). A second problem is the issue of contradicting appraisals, e.g., (o, d, r) and (o, d ,$-r$) at the same time. What should be done with both elements? Concerning these issues we have chosen to force specific theories to think about this by never allowing two elements that only differ in terms of their intensities in the sets V or I. This explains the additional constraint defined in Definitions 2.3 and 3.3. Depending on the specific appraisal theory being formalized, this could be resolved by, e.g., defining an additional constraint such that whenever this would occur, values are summed into a new element. By doing so, contradicting appraisals are also resolved: with this specific constraint they cancel each other. We do not enforce this specific resolution, but our formalism enforces a theory to think about the intensity issue (see also; Reilly, 2006).

Suppose that our baby has three emotions: calm, distressed and neutral: $E=\{calm,dis,neut\}$. The set PI is filled with all possible combination for (e,r) tuples. The baby has one mediating process $M=\{m_e\}$ that relates PV (the set of potential appraisal dimension values) to PI (the set of potential emotion-component intensities) in the following way: $m_e(\{(po_d,pleas,-1),(po_d,sudd,1)\})=\{(calm,0),(dis,1),(neut,0)\}$. 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. Again, I equals $\{(calm,0),(dis,1),(neut,0)\}$ as $V=\{(po_d,pleas,-1),(po_d,sudd,1)\}$. The baby is distressed as a result of appraising a noisy dog.

3.4 Appraisal- theory specific dependency between processes

Our formalism represents processes connected to each other via different kinds of guarded dependencies. Many appraisal theories explicitly define relations between processes. For example, simpler appraisal processes such as suddenness and intrinsic pleasantness that specifically trigger more complex processes such as urgency (Scherer, 2001). To be able to define such relations we introduce guards and links.

Definition 4.1: The set $G=\{g_1,\ldots,g_n\}$ of guards is a set of second-order predicates over the elements of the sets introduced in the previous definitions. This allows the definition of conditional dependencies between processes.

Definition 4.2: Let L be the set $L=\{l_1,\ldots,l_n\}$ with $L\subseteq PP\times PP\times G$ 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), with $p,q\in PP$, $g\in G$, processing in q is influenced only if the guard g is true and process p is active itself[4]. Process dependencies can be of different types, such as excitatory or inhibitory.

---

[4] Note that when a structural theory mentions dependencies between processes without mentioning any activation conditions, G can be defined as $G=\{true\}$, such that all dependencies have guard $g=true$ 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.

When a specific theory needs dependency types, this is represented by defining subsets of L, e.g., $INHIBITORY \subseteq L$.

For our baby, suppose there are four dependencies $L=\{l_1, l_2, l_3, l_4\}$ between the perception, appraisal and emotion generation processes:

- $l_1=(p_h, a_p, (\exists x\ x \in O))$, i.e., if an object is present in O, appraisal process a_p is activated by perception process p_h,
- $l_2=(p_h, a_s, (\exists x\ x \in O))$, i.e., the same for process a_s,
- $l_4=(a_p, m_e, (\exists x\ x \in V \wedge x=(d,i) \wedge i<>0))$, $d \in D$, i.e., m_e is activated by a_p when something is appraised,
- $l_3=(a_s, m_e, (\exists x\ x \in V \wedge x=(d,i) \wedge i<>0))$, $d \in D$, i.e., same for a_s.

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

3.5 Appraisal- theory specific 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, as explained earlier, 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, unlike the guards used to represent activation conditions.

Definition 5.1: The set $H=\{h_1, \ldots, h_n\}$ of guards is a set of second-order predicates over the elements of the previously introduced sets (except G and L).

3.6 What does our 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 she 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 conductivity 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 conductivity 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.

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 (*S*cherer, *S*mith and *K*irby). 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 paragraph 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.

4.1 Scherer's SEC model and Smith and Kirby's ADM

Scherer's SEC model. This model is based on the idea that appraisal processes evaluate stimuli in a certain sequence (for simplicity, in this article 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 1. For the current integration we restrict ourselves by excluding norm/self compatibility.

Table 1.

Overview of the stimulus checks related to novelty, pleasantness and coping potential existing at the sensory-motor, schematic and conceptual level (Scherer, 2001). For explanation see text.

	Novelty	**Pleasantness**	**Goal/Need conduciveness**	**Coping potential**
Sensory-Motor level (innate)	Sudden, intense stimulation	Innate preferences/ aversions	Basic needs	Available energy
Schematic level (automatic)	Familiarity: schema matching	Learned preferences or aversions	Acquired needs, motives	Body schema (automated knowledge of how the body

				functions, what it can do etc.)
Conceptual level (deliberative)	Expectations: cause/effect, probability	Recalled, anticipated, or derived positive-negative estimates	Conscious goals, plans	Problem-solving ability

In general, sensory-motor level appraisal processes are related to biological needs and drives and to biological mechanisms, and are mostly genetically determined. As an example, consider chocolate. Chocolate is fat and contains sugar. Eating it makes us happy, not because of some deep cognitive evolution, but because we are genetically tuned to like fat and sugar because it helped survival. Schematic-level appraisals are based on learned knowledge organized into schemas. As an example consider watching your favorite television show. Over time, this event has become a kind of end-of-day goal. It makes you comfortable when achieved, and frustrated when not. However, there neither is an innate preference for this, nor any long-term conscious contemplation about why this is or isn't good for you. Conceptual level appraisal processes are based on propositional-symbolic, cortical mechanisms that require consciousness (Scherer, 2001). As an example consider failing an important exam. This triggers reasoning efforts about long-term job opportunities, about whether or not one expected it to happen based on prior exams, about related exams in the future, about whether one could have done better, etc. In general, higher levels of processing are used to appraise the situation if lower levels seem inadequate to evaluate the stimuli. For example, if you are a strong muscled man and you witness an armed robbery, your first reaction might be to run (in slight anger) towards the robber and try to overpower him/her (a sensory-motor/schematic appraisal). However, as one realizes that the weapon might be used against yourself, and that your partner and kids might greatly suffer in that case, you opt for immediately calling the police while keeping an eye on the situation (conscious contemplation). You reevaluated the situation in a different way, because of its complexity.

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).

- 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, evaluated continuously and operate on different levels of processing, they are sequential in the sense that the later steps are only deployed to the maximum if earlier checks indicate that this is necessary. Furthermore, later steps, in general, need more elaborate processing. 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 (2001). 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 toward 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.

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 and 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 pleasantness, and other biologically important survival information. In contrast, the latter two modes are considered to be cognitive modes, and are considered more important. 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 and Kirby, 2001). 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.

Short comparison of both models. The ADM assumes three modes of information processing (stimulus perception, associative processing and reasoning). These modes generate appraisal information. 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. The SEC model is thus more specific about the actual appraisals that can take place, and on which level. However, both models assume three of these processing levels on which appraisal takes place. A slight difference regarding these levels is that the SEC model emphasizes the sensory-motor level as a full-blown contributor to appraisal, while the ADM assumes that “most of the information processed by the appraisal detectors is generated through either associative processing or reasoning”.

The appraisal information from the processing modes in the ADM model 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. In the SEC model this appraisal state also exists in the form of appraisal registers (accumulating appraisal information and detecting preliminary closure). However, the SEC model provides more detail into why such integration should exist, apart from its function related to feeding back information to appraisal processes and mediating between appraisal and emotion. The SEC model explicitly argues that this appraisal state mediates between the activation of different appraisal steps (i.e., first novelty and pleasantness, then goal conduciveness, etc.). These checks evaluate the stimuli in a specific order and depend on one another, a feature not explicitly represented in the ADM.

A key difference that is also reflected in the formal specification presented next is that the SEC model specifically defines relations between level of information-process and individual stimulus checks. For example, goal/need relevance takes information from all three levels of processing, while intrinsic pleasantness only from the sensor-motor level.

Anther key difference is that the ADM does not explicitly define "information producing processes" (perception, contemplation etc.) *and* appraisal processes, while the SEC model does. In the SEC model stimuli are perceived and then appraised by a stimulus check, in the ADM this is all done in one step by the *processing mode*.

4.2 Formal integration, the SSK model

In this section we present and explain 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. The specification is also presented in graphical form in Figure 3. This figure is intended as an overview of the formal description. The actual set notation and its explanation is detailed in the rest of this section and in Appendix A, in which a simplified version of the SSK model (used for the PacMan game experiment, Section 5) is fully specified.

(figure 3 about here)

Before describing the integrated model, some terminology has to be introduced. 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 $p \in P$ in our formal notation. This is counterintuitive, as a perception process in normal terminology would be responsible only for mapping sensory information to internal representations. However, as our formalism uses a different meaning to perception process, we can explain this as follows. A perception process is the only process responsible for generating to-be-appraised objects in working memory. As such, any mode of information processing that generates (or needs) to-be-appraised objects, such as schematic reasoning and related appraisal processes, must be a formalized by at least one perception process that *can* "generate" these objects. The fact that in the ADM the mode of reasoning not only generates to-be-appraised objects but actually *appraises* objects is irrelevant: our formalism is able to express more detail as it has a different construct for appraisal processes, which is why we should separate appraisal from object-generation. In the SEC model the level of appraisal refers to the dominant type of information manipulation involved in appraising a situation. For example, appraisal processes at the schematic level mostly use schematic information, and not, e.g., sensory-motor information. Again, to produce objects resulting from different types of information processing levels, we need to introduce these levels as perception processes. However, it should be noted that the level of appraisal (as in the SEC model) is not *limited to* the perception process; the level extends to all appraisal processes that receive information from the sensory-motor perception process. We refer to these perception processes as levels of appraisal, because they are the necessary prerequisites for these levels, and because they activate appraisal processes residing in these levels. As a whole, we feel that this way of formalization makes explicit what kind of information processing is needed for what appraisal process. As a result, we can clearly see, e.g., in Figures 3a-c, which appraisal processes are influenced by what *kind of information*, and therefore we know at what level of complexity such an appraisal process must be able to function.

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 choose to focus on processes (perception, appraisal and mediation) and their dependencies.

Formalizing 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}`

Formalizing 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. As we have argued above, this is consistent with both models. The set `P` is represented by the white boxes in Figure 1 and equals:

`P={stimulus_perception, schematic, 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, i.e., there is a subset `EXCITATORY⊆L` containing these dependencies. 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 Appendix A.

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.

Formalizing 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={{the elements of the set of all stimulus checks in the first
three steps of the SEC model}, 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.

Formalizing mediating processes. The three rightmost colored boxes represent mediating processes. The set `M` contains the following elements:

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

These mediating processes are positioned between the different appraisal steps. Mediating processes are activated by the appraisal processes of one appraisal step and activate appraisal processes of another appraisal step, 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 didn't 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`:

$$(\exists x\ x\in V \wedge x=(o,d,r) \wedge r>t \wedge d=\texttt{novelty\_dim}),\ \text{with}\ \texttt{novelty\_dim}\in D\ \text{and}\\ t\in[0,1]\ \text{an arbitrary threshold.}$$

This guard checks the existence of a `novelty_dim` value greater than an arbitrary threshold `t`. Only if this value exists, the will be guard true, and thus the connection is active. Now the `novelty` appraisal process excites the `relevance` detector.

Formalizing 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` (i.e., a second subset $\mathtt{INFORMATION\_AVAILABLE} \subseteq \mathtt{L}$ containing these links is defined). This means that when the guard of the dependency is true, the target process is informed of the fact that new information is available.

Formalizing the "actual" SEC levels of appraisal and appraisal steps.

Our formalism enables us to define subsets of processes, such that processes can be grouped. This enables us to formally describe the levels of appraisal as defined in the SEC model. As an example we formalize the sensory-motor level. It is defined as $\mathtt{A}_{SM} \subseteq \mathtt{A}$, with

$$\mathtt{A}_{SM}=\{\mathtt{suddenness,novelty,intrinsic\_pleasantness,goal/need\_relevance,goal/need\_conduciveness,urgency,power}\}$$

Note that the processes involved are exactly those that appraise based on sensory-motor information. Also note that the processes involved in sensory-motor based appraisal are not the same as those involved in the first appraisal step. For example, `familiarity` is included in the first step, but not in the first level. This is normal, as steps (and the appraisal processes therein) can appraise on multiple levels, and levels can contain appraisal processes from different steps. This reflects the parallelism inherent in the SEC and ADM models. Every level of processing can activate appraisals, and many appraisals can concurrently be active, even at different levels.

Appraisal steps can be formalized as different subsets of `A` and `M`. For example, the first step related to relevance detection can be a subset $\mathtt{A}_R \subseteq \mathtt{A}$, with all green appraisal processes (Figure 3) as elements, while $\mathtt{M}_R \subseteq \mathtt{M}$ would contain the mediating process `relevance_detector`.

4.3 Summary

Integration and comparison are important reasons to formalize appraisal theories (Wehrle and 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 analysis and 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.

Further, it is now possible to ask formal questions about the model, for example about the intersection of steps and level subsets. When implementing the formal description, using e.g. a Prolog-like language, questions such as “give me all processes that are (level-2 or level-3) *but not* (step-2 *or* step-3)” can be answered automatically. If such processes would exist, this would mean they appraise using higher-level knowledge, but are not specifically implicated in complex situations. These processes would be the ones that are predicted to be effortful *and* often used (as step 1 is used first). For goal/need relevance checking this seems to be the case: it *can* be quite effortful and is often needed to evaluate the relevance of a stimulus on a higher (e.g., deliberative) level. Such processes are of interest to the SEC model, as it assumes that earlier steps generally are less effortful. It is also interesting (though not entirely surprisingly, as the SEC model is an appraisal theory) to see that goal-related processing is a cornerstone of the SEC model: it uses information from all three levels, and is implicated in the first two steps, most heavily used steps. What is also nice to see is that the SEC and ADM models are quite compatible. This provides strong support for their individual credibility (assuming our integration is not entirely nonsensical, of course).

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 Java 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 analyze the behavior of the emotional PacMan agent, specifically related to the issue of appraisal-intensity integration.

Why PacMan? PacMan-like environments have been used to research emotion-related phenomena, both in the appraisal-theoretic domain (Wehrle and Scherer, 2001) as well as the virtual agent domain (Broekens and DeGroot, 2004). PacMan (Figure 5) is 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.

5.1 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, that the formal notation presented in this article is a powerful one. Consequently, the goal of this experiment was not to design a believable or "full-blown" emotional agent.

Simplifications. 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 defending 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. Importantly, 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 4. The full set notation can be found in Appendix A.

(Figure 4 and 5 about here)

Computational model: filling in missing declarative information. To be able to build an executable model, we need to address several issues mentioned earlier, 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 provides formal support for this view. 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 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 `goal/need_relevance`. Every outgoing dependency from the `relevance` detector to an appraisal process of the next appraisal step has a guard equal to:

$$
\begin{gathered}
(\exists v_1, v_2, v_3\ \ v_1, v_2, v_3 \in V \wedge v_1 = (o, d_1, r_1) \wedge v_2 = (o, d_2, r_2) \wedge v_3 = (o, d_3, r_3) \wedge \\
(|r_1| + |r_2| + |r_3|)/3 > 0.15 \wedge d_1 = \texttt{novelty\_dim} \wedge d_2 = \texttt{intrinsic\_dim} \wedge \\
d_3 = \texttt{relevance\_dim})
\end{gathered}
$$

We assume all three tuples v_1, v_2 and v_3 to exist. If not, we take their corresponding activation value to be 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 appraisal step has a guard equal to:

$$(\exists v_1, v_2 \; v_1, v_2 \in V \wedge v_1 = (o, d_1, r_1) \wedge v_2 = (o, d_2, r_2) \wedge (|r_1| + |r_2|)/2 > 0.25 \wedge$$
$$d_1 = \text{conduciveness\_dim} \wedge d_2 = \text{urgency\_dim})$$

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 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:

$$(\exists v_1, v_2 \; v_1, v_2 \in V \wedge v_1 = (o, d_1, r_1) \wedge v_2 = (o, d_2, r_2) \wedge r_1 * r_2 > 0 \wedge d_x = \text{control\_dim} \wedge$$
$$d_y = \text{power\_dim})$$

Again we assume that both tuples $\mathtt{v}_1$ and $\mathtt{v}_2$ exists, and if one of them (or both) do not, we take their corresponding value to be 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 pleasurable 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 2). The schematic perception process reacts to events that relate to the goal of collecting points (Table 3).

Table 2.

PacMan appraisal related to survival need

Appraisal process	**Dimension**	**Checking criteria**
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 PacMan has and the amount of lives left (hunger is simulated based on the last time PacMan 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 3.

PacMan appraisal related to the goal of gathering points.

Appraisal process	**Dimension**	**Checking criteria**
Familiarity	`novelty_dim`	Seeing a dot is more common than a ghost is, and seeing a ghost is more common than a power-pill which again is more common than fruit is.
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.

5. 2 Experimental setup and verification of the computational model

We have instrumented PacMan based on the just described simplification of the SSK model. We simulated perception (i.e., the generation of to-be-appraised objects) with a simple decision support system that generates mental objects based on PacMan's current perception of its direct surroundings (of a radius of 4 positions in the maze grid). 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 simplified SSK specification. These appraisal processes produce appraisal dimension values, as specified in Tables 2 and 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 Appendix A).

We first implemented integration simply by adding appraisal values that belong to the same appraisal dimension and storing the result in the appraisal state (so we assume a naive integration of the intensity of different appraisal results). This is an important fact for the analysis of the observed appraisal behavior of the agent.

The experiment itself consists of one trial in which a human player who plays the first level of the PacMan game (by eating all dots), loses a life two times during that level, and eats several ghosts. While playing, the activity of the levels of appraisal is logged, i.e., `relevance`, `implication` and `coping_potential`. These activities are plotted over time in Figure 6. 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 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). This must mean that the intensity of the appraisal-dimension values that feed to the level of `implication` is not high enough to trigger the next appraisal step.

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 strongly relates to the appraisal intensity issue already discussed earlier, and our naive assumption that intensities can simply be added. The `conduciveness_dim` dimension is bi-polar 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-conducive 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. As a potential solution, we introduced one extra appraisal dimension, an extra appraisal process that checks stimuli related to `non-conduciveness`, and a link between that process and `implication`. When rerunning the experiment, the new results were as expected. Coping potential *is* activated (and therefore we know implication is high enough to activate the next level of appraisal) but low since PacMan has not eaten a power-pill recently (Figure 6b, between $9000 \leq t \leq 13000$). This can be easily explained: the addition of the new appraisal dimension and associated elements separates between positive and negative conduciveness. As a result, the computational model can now naively integrate appraisal intensities (positive conduciveness does no longer cancel out negative conduciveness), and the positive and negative conduciveness together are strong enough to activating the next appraisal step.

5.3 Summary

Our formalism helped to develop a computational model based on a combination of two theories of appraisal, 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, and propose a potential solution to, what was in our case a problem related to appraisal intensity. Note that we do not claim anything about our particular solution for this intensity problem; we claim that the formal description helped us to develop and debug our computational model, i.e., to detect this intensity problem.

(Figure 6a and 6b about here)

6 Discussion

We first discuss additional appraisal-related phenomena in the context of our formalism, then we address several formalization issues. Lastly, we discuss related and future work.

6.1 Integration of appraisals, emotion-appraisal-cognition interplay, structure vs. process.

In this section we address several appraisal related phenomena in the context of our formalism. We show how the formalism can handle (or be extended to handle, or can not handle) these phenomena. We end this section with some words on the relation between process- and structural models.

Our formalism enables the formal description of the influence of emotion on appraisal. An example of how the formalism handles this is now shown. Again for the baby example, consider the mental object to be appraised $\mathtt{po}_{\mathrm{dog}} \in \mathtt{O}$, an emotion component representing the feeling of the baby (as in our example) $\mathtt{e}_{\mathrm{happy}} \in \mathtt{E}$, an appraisal dimension $\mathtt{d}_{\mathrm{suddenness}}$, and an appraisal process that checks for suddenness $\mathtt{a}_{\mathrm{suddeness}} \in \mathtt{A}$, with $\mathtt{a}_{\mathrm{suddenness}}(\{\mathtt{po}_{\mathrm{dog}}, (\mathtt{e}_{\mathrm{happy}}, 1)\}) = \{\mathtt{po}_{\mathrm{dog}}, \mathtt{d}_{\mathrm{suddenness}}, 0.5\}$, and $\mathtt{a}_{\mathrm{suddenness}}(\{\mathtt{po}_{\mathrm{dog}}, (\mathtt{e}_{\mathrm{happy}}, -1)\}) = (\mathtt{po}_{\mathrm{dog}}, \mathtt{d}_{\mathrm{suddenness}}, 1\}$. This description formally models a baby who appraises the dog as very sudden if the baby is unhappy at that moment, while it appraises the dog as moderately sudden when it is happy. Depending on the value of the emotion component e_{happy} the baby appraises the situation differently.

As already mentioned in the definitional part of this article, the formalism can be extended with new concepts using the standard set-theoretic paradigm. For example, should a theory need different types of objects, such as beliefs, desires and intensions, or motives and drives for that matter, subsets of the PO of available objects can be defined (such as $MOTIVES \subseteq PO$). Data and dependency constraints can now be defined over these subsets instead of over the set PO. For example, if our baby is hungry and thirsty, we might formalize these as o_{hungry}, $o_{thirsty} \in DRIVES,\ DRIVES \subseteq PO$. Another example of how the formalism might be extended is by including time. This would open up a complete new realm of modeling, including formalizing the detailed (over time) course of emotion elicitation, the bodily reactions, etc. It would also enable to proof that some event had to be preceded by another, instead of just being present as in the current formalism. Time can be included into the formalism by defining timed-slices of sets, a sort of frames. Subsets of sets are defined, with a time index. The definitions presented can be reformed to include the time index. Work is underway into this direction, although preliminary results show that time does render the formalism significantly more complex.

Regarding the interplay between cognition and emotion, we have to first mention that the way we have modeled the relations between the different SECs is very simplified. More complex relations then excitatory ones are needed to capture these relations. As an example of a more complex relation between emotion and cognition, not necessary related to the SEC model, consider the following. We assume a perception process $p_{stimulus\_oriented} \in P$, and a perception process $p_{topdown} \in P$. We also assume an emotion component $e_{valence} \in E$, and a mediating process $m_{valence} \in M$. It is known that positive valence is related to top-down, integrative processing, while negative valence is related to stimulus-oriented, critical processing. We define the following dependency relations to formalize this: l_1, l_2, l_3, $l_4 \subseteq L$, such that $l_1, l_3 \in EXCITATORY$, $l_2, l_4 \in INHIBITORY$, $l_1 = (e_{valence}, p_{top\text{-}down}, \exists i \in I$ with $i = (e_{valence}, r)$ and $r > 0)$, $l_2 = (e_{valence},\ p_{top\text{-}down},\ \exists i \in I$ with $i = (e_{valence},\ r),\ r < 0)$, $l_3 = (e_{valence}, p_{stimulus\_oriented}, \exists i \in I$ with $i = (e_{valence}, r)$ and $r < 0)$, $l_4 = (e_{valence},\ p_{stimulus\_oriented},$

$\exists i \in I$ with $i=(e_{valence}, r), r>0)$. These dependencies define the relation between top-down and bottom-up processing and valence: positive valence increases the tendency to process incoming information in a top-down, integrative way, while negative valence decreases this tendency. Note that ideally we would need a notion of process activity here, to describe *how* active a process actually is. This is currently not represented in our formalism, but is needed once these process-oriented descriptions need to be formalized. So, although some of the cognition-emotion relations might be captured with our formalism, the level of detail is limited to structural relations, and consequently there exist relations not describable with the current version of our formalism.

Our formalism focuses on structural models, e.g., it currently does not include a concept of time and it does not formalize dynamic relations or equilibriums between processes. However, regarding the relation between process and structure, we have to make a reflecting remark. There is a finer line between structural and process models than we have drawn in the beginning of this article. A structural model, describing, e.g., appraisal dimensions, can be augmented with process dynamics to describe how these processes interact in detail. Consider for example the work by Scherer (2001). Although parts of his theory are definitely structural (listing of appraisal processes, the relation between processes and emotion components), some are definitely process-oriented (sequential evaluation of appraisal steps, integration of appraisal results in appraisal registers). In general we could say that the detail increases from structural to process (as argued in Broekens & DeGroot, 2006), however, difference in detail alone is not enough to differentiate between both types of models. A better one is the difference in kind of representation used in structure and process models, and therefore the kind of things expressible by the model. For example, any process needs a notion of time, so any process model of emotion elicitation is forced to say something about the course of that elicitation, the sequence of evaluations, the intensity over time, etc. Emotional coping, for instance, is a phenomenon that is difficult to model using *only* structural models, specifically as the type of coping and the effectiveness of coping is dependent on the time course of the emotion to-be-coped with. Consider the following example. A child might perfectly cope cognitively with its feelings of anger due to bullying (e.g., they are the ones that are stupid, they are not nice because they are actually unhappy that's why they pick on me, etc.), while have extreme difficulty suppressing anger reactions when put in a mildly unpleasant situation by surprise (in which case a more direct, emotion oriented coping process would be involved). Expressing these two ways of coping with a purely structural model will be difficult, as time and emotion "build-up" plays an important role.

It might seem contradictory that we have just argued that coping is in need of process-orientation, while our formalism is structure oriented and we have modeled coping processes and argued in Section 3 that some elements of coping can be represented with our formalism. This contradiction is easily explained by the fact that our formalism indeed contains several process-oriented elements, e.g., intensity and conditional process dependencies. Without these, it would be impossible to start talking about phenomena such as coping. However, the formalism itself is not specifically well suited to represent process-oriented theories in full glory, at least not without the concept of time and process activity.

6.2 Issues with formalization.

We now discuss several issues with formalization per se. 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 1987). 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, since formal specifications have many benefits, the formal specification will function as a substitute for the theory. This could result in overly strict interpretations of the 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.

Third, using a set-based formalism introduces the need for careful translation from written theories to formal descriptions. Much detail can be involved, especially when relations between processes are being formalized. This can, in some cases, result in large descriptions. One might conclude that formal descriptions are more difficult to read and debug than a “normal” computational model written in an imperative language. In general however, the question is how to select the right level of abstraction for the formal description. Once this level has been found, only several lines of “formalism code” need to be written to specify an appraisal structure. However, we have to keep in mind that a formal description is not supposed to do anything. It is a *specification* used to check experimental outcomes, simulated outcomes etc. in relation to a theory. So, the formal description of the theory is at a higher level than any “runnable” program would be. For example, our formalism allows to notate things like “if there is an appraisal process active related to the appraisal dimension goal_conduciveness, there must be at least one perceived goal and an event that either helps that goal or not”. This is one line in our notation, and covers a generic assumption of appraisal theory. This line can then be used to check the consistency of a running program. As such, the higher expressive power of the set-based notation and the higher level of abstraction used enable the description of certain relations in a very concise way. Further, one is never forced to formalize a complete theory. The formalism can very well be used to define *parts of* a theory. The resulting formal description can then still be used to check parts of the theory against (simulated) experimental outcomes.

We have, for example, implemented an extended version of the baby example in Prolog (a logic-based programming language), this resulted in only half a page of prolog code, and it enabled us to ask formal questions about perceived objects, active appraisals, resulting emotion components, etc. A “normal” program cannot be asked such questions. On the other hand, a “normal” program would be able to express more detail, especially regarding the processes involved in appraisal. We have to keep in mind that our formalism is targeted at the structure of appraisal. Adding to much process-oriented detail into a formal description of a theory (or experiment for that matter) would indeed render the formalism a difficult one to use. However, as shown in the discussion, it is quite straightforward an exercise to model non-trivial relations between emotion and cognition, as well as appraisal influenced by emotion. The ability to use a standard notation to describe such relations at a high level is, in our opinion, a useful addition to the current state of affairs.

Of course there are caveats. One would have to get accustomed to set-notation and Prolog like code in order to efficiently use our formalism, and, there will certainly be situations in which our formalism is not a handy tool, e.g., when processes must be modeled in detail. In such case we would suggest using other models, such as the one proposed by Gratch and Marsella (2004).

The last issue are actually relates to the “usability” of the formalism itself. Several additional remarks have to be made regarding this. First, although it seems we have abandoned the formal notation when comparing the SEC and ADM models, this is not true. The graphical representation is a result of using the formal constructs (see, e.g., the issue with perception processes $p \in P$ and levels of appraisal, as discussed in Section 4.2). The elements used in the graphical representation directly map onto the formal description of the SSK model. Further, the formal notation can be used to proof statements about theories and experimental data, while the graphical notation is used to give insight into the structure of the model. The fact that, for explanatory purposes, we prefer a graphical form, does not do away with the utility of a common formal notation to describe appraisal theories at an abstract level. Compare, e.g., the field of software engineering (see also, Broekens & DeGroot (2006) on this issue). The fact that a common, high-level, abstract, notational scheme such as UML is used to formally describe software architectures, does not exclude the use of a graphical representation of such a formal description. Doing so would be a disaster for the understandability of the notation. Also, the fact that such a high-level description cannot fully describe all the functionally needed to “program” the system (analogous to a formal description of a theory not being detailed enough to program the complete emotional agent) does not do away with the usefulness of having an abstract description in a common formal language.

Second, the models we have integrated (SEC and ADM) are quite compatible. It would, perhaps, be stronger evidence for the usefulness of the formalism to compare two *very different* theories. However, we feel this is an issue of choice. If these theories are very similar, then a potential integration between them is useful per se, and our formalism greatly helped us doing so. We have chosen to integrate two theories as an example of the use of the formalism. Would we have chosen to compare two theories, a comparison between two more different ones would have been more interesting, e.g., a comparison between the OCC model (Ortony et al., 1988) and the SEC model (Scherer, 2001). However, doing both in this, already lengthy, article, we feel is not needed. We consider such comparison critical, though future work.

To conclude, we are aware that certain levels of descriptive detail are unattainable, and in some cases undesirable, and we know that other uses of our formalism must be investigated. However, this is not a problem for the validity and usefulness of the formalism we propose now.

6.3 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. We also relate our formalization approach to the study of Individual-Based models.

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 and 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 we would introduce time (see future work) in our formalism it would allow 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.

Although not directly related to our formal approach, several computational models of emotion elicitation and appraisal are certainly relevant to what we have done, and should be mentioned. For example, the work by Elliott (1992) presents a vast expansion and computational implementation of the cognitive appraisal model by Ortony and others (1998). The work by Gratch and Marsella (2004; but also a lot of earlier work) presents a generic framework to experiment with (and use what is known about) cognitive theories of emotion, specifically related to the influence of planning, coping and decision-making. Finally the work by Hudlicka (2004) presents a computational framework specifically aimed at investigating the nature and degree of influence of different emotional personalities on cognitive processing and actual behavior of agents. The main difference between these models and ours is that our formalism is a tool to formally specify, in detail, structural relations in appraisal theory, while these models are tools that aim at understanding the processes (specific object evaluations, influence of emotion and personality on behavior and cognition over time, etc.) underlying emotion by means of a computational model.

6.4 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 et al. 2001) or coping (Lazarus, 2001), without the LTM construct.

Finally, a comparison between the structuralist approach towards theory formalization and our approach is planned, as well as an attempt to formally compare the OCC model (Ortony et al., 1998) with the SEC model (Scherer, 2001).

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 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 useful 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 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 facilitates development and evaluation of computational models of emotion based on cognitive appraisal theory. Although not shown in this paper, our formalism can, in principle, also be used to formalize less cognitively-oriented theories.

Acknowledgments

We thank two anonymous reviewers for their excellent criticism. Without exception, all of their remarks have considerably enhanced the readability, consistency and completeness of our formalism, and often provided us with new insights.

References

Baars, B. J. & Franklin, S. (2003). How conscious experience and working memory interact. *TICS, 7*(4), 166-172.

Breazeal, C. (2001). Affective interaction between humans and robots. In: J. Keleman and P. Sosik (eds.), *Proc. of the ECAL 2001, LNAI 2159* (pp. 582-591).

Broekens, J. & DeGroot, D. (2004). Scalable and flexible appraisal models for virtual agents. In: Q. Mehdi and N. Gough (eds.), *Proc. of the 5th Game-On International Conference* (pp. 208-215).

Broekens, J. & DeGroot, D. (2006). Formalizing cognitive appraisal: From theory to computation. In: R. Trappl (ed.), *Proc. of the 19th European Meeting on Cybernetics and Systems Research* (pp. 595-600).

Chow, B. (2003). *PacMan*. http://www.bennychow.com/pacman_redirect.shtml, Retrieved: 1-Nov-2003.

Coddington, A. & Luck, M. (2003). Towards motivation-based plan evaluation. In: I. Russell and S. Haller (eds.), *Proc. of the Sixteenth International FLAIRS Conference* (pp. 298-302).

Elliott, C. D. (1992). *The affective reasoner: a process model of emotions in a multi-agent system*. PhD Thesis, Northwestern University, Computer Science.

Franklin, S. & Graesser, A. (1999). A software agent model of consciousness. *Consciousness and Cognition, 8*(3), 285-301.

Frijda, N. H. & Mesquita, B. (2000). Beliefs through emotions. In: N. H. Frijda, A. S. R. Manstead and S. Bem (eds.), *Emotions and Beliefs* (pp. 45-77), Cambridge University Press.

Frijda, N. H. & Zeelenberg, M. (2001). Appraisal: what is dependent? In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 141-156), Oxford University Press.

Gmytrasiewicz, P. & Lisetti, C. (2002). Emotions and personality in agent design and modeling. In: J.-J.Ch. Meyer and M. Tambe (eds.), *Proc. of Intelligent Agents VIII, LNAI 2333* (pp. 21-31).

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

Heylen, D., Nijholt, A., Akker, R. op den. & Vissers, M. (2003). Socially intelligent tutor agents. In: T. Rist, R. Aylett, D. Ballin and J. Rickel (eds.), *Proc. of the 4th International Workshop on Intelligent Virtual Agents (IVA 2003)* (pp. 341-347).

Hudlicka, E. (2004). Beyond Cognition: Modeling Emotion in Cognitive Architectures. In: *Proc. of the Sixth International Conference on Cognitive Modeling (ICCM04)* (pp. 118-123).

IEEE (2000), *IEEE Recommended Practice for Architectural Description of Software-Intensive Systems*, Std 1471-2000, IEEE.

Jennings, N., Sycara, K. & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems, 1*(1), 7-38.

Lazarus, R. S. (2001) Relational meaning and discrete emotions. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 37-67), Oxford University Press.

Levine, L. J., Prohaska, V., Burgess, S. L., Rice, J. A. & Laulhere, T. M. (2001). Remembering past emotions: The role of current appraisals. *Cognition and Emotion, 15*(4), 393-417.

Lewis, M. D. (2001). Personal pathways in the development of appraisal. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 205-220), Oxford University Press.

MacNamee, B. & Cunningham, P. (2003). Creating socially interactive non player characters: The μ-SIC system. *IJIGS, 2*(1), 28-35.

Marsella, S. & Gratch, J. (2001). Modeling the interplay of emotions and plans in multi-agent simulations. *Proc. of the 23rd Annual Conference of the Cognitive Science Society* (pp. 294-299).

Meyer, J.-J.Ch. (2004). Reasoning about Emotional Agents. In: R. López de Mántaras and L. Saitta (eds.), *Proc. of the 16th European Conference on Artificial Intelligence* (pp. 129-133).

Oatley, K. (1999). Emotions. In: R.A. Wilson and F. Kiel (eds.), *The MIT Encyclopaedia of the Cognitive Sciences*, MIT Press (pp. 273-275).

Ortony, A., Clore G. L. & Collins A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press.

Partala T. & Surakka, V. (2004). The effects of affective interventions in human-computer interaction. *Interacting with Computers, 16*, 295-309.

Picard, R. (1997). *Affective Computing*. The MIT Press.

Reekum, C. van. (2000). *Levels of Processing in Appraisal: Evidence from Computer Game Generated Emotions*. PhD Thesis nr. 289, University de Geneve, Section de Psychology.

Reilly, W. S. N. (2006). Modeling what happens between emotional antecedents and emotional consequents. In: R. Trappl (ed.), *Proc. of the 19th European Meeting on Cybernetics and Systems Research* (pp. 607-612).

Reisenzein, R. (2000). Wundt's three-dimensional theory of emotion. In: W. Balzer, J. D. Sneed and C. U. Moulines (eds.), *Poznan Studies in the Philosophy of the Sciences and the Humanities* (vol. 22, pp. 219-250).

Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 187-204), Oxford University Press.

Roseman, I. J. & Smith, C. A. (2001). Appraisal theories: Overview, assumptions, varieties, controversies. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 3-19), Oxford University Press.

Scherer, K. R. (2001) Appraisal Considered as a Process of Multilevel Sequential Checking. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 92-120), Oxford University Press.

Smith, C. A. & Kirby, L. D. (2000). Consequences require antecedents: Toward a process model of emotion elicitation. In: J. P. Forgas (ed.), *Feeling and Thinking: The Role of Affect in Social Cognition* (pp. 83-108), Cambridge University Press.

Wehrle, T. & Scherer, K. R. (2001). Towards computational modeling of appraisal theories. In: K. R. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: Theory, Methods, Research* (pp. 350-368), Oxford University Press.

Westmeyer, H. (1989). Psychological theories from a structuralist point of view: A first introduction. In: H. Westmeyer (ed.), *Psychological Theories from a Structuralist Point of View* (pp. 1-12), Berlin: Springer.

Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In: J. P. Forgas (ed.), *Feeling and Thinking: The Role of Affect in Social Cognition* (pp. 31-58), Cambridge University Press.

Appendix A: Formal Appraisal Structure of the Simplified SSK model applied to PacMan

Perception appraisal and mediating processes (just the processes, not the formal description of their input-output relations):

```
P={  stimulus_perception, schematic}
A={  suddenness, familiarity, novelty, intrinsic_pleasantness,
     relevance, conduciveness, urgency, control, power}
M={  relevance_detector,
     implication_detector, coping_potential_detector}
```

Mental objects, appraisal dimensions and emotion components:

```
PO={ see_ghost, lost_ghost, eaten_by_ghost, see_edible_ghost,
     lost_edible_ghost, belief, eaten_ghost, see_power, eaten_power,
```

see_dot, eaten_dot, see_fruit, lost_fruit, eaten_fruit}

D={ novelty_dim, intrinsic_pleasantness_dim, conduciveness_dim, relevance_dim, urgency_dim, control_dim, power_dim}

E={}

Guards and dependencies:

G={ true, guard1, guard2, guard3} with:

guard1=($\exists v_1, v_2, v_3 \quad v_1, v_2, v_3 \in V \wedge v_1=(o, d_1, r_1) \wedge v_2=(o, d_2, r_2) \wedge v_3=(o, d_3, r_3) \wedge (|r_1|+|r_2|+|r_3|)/3>0.15 \wedge d_1=\text{novelty\_dim} \wedge d_2=\text{intrinsic\_dim} \wedge d_3=\text{relevance\_dim}$)

guard2=($\exists v_1, v_2 \quad v_1, v_2 \in V \wedge v_1=(o, d_1, r_1) \wedge v_2=(o, d_2, r_2) \wedge (|r_1|+|r_2|)/2>0.25 \wedge d_1=\text{conduciveness\_dim} \wedge d_2=\text{urgency\_dim}$)

guard3=($\exists v_1, v_2 \quad v_1, v_2 \in V \wedge v_1=(o, d_1, r_1) \wedge v_2=(o, d_2, r_2) \wedge r_1 * r_2>0 \wedge d_x=\text{control\_dim} \wedge d_y=\text{power\_dim}$)

L={ (stimulus_perception, suddenness, true),
(stimulus_perception, intrinsic_pleasantness, true),
(stimulus_perception, relevance, true),
(stimulus_perception, conduciveness, true),
(stimulus_perception, urgency, true),
(stimulus_perception, power, true),
(schematic, familiarity, true),
(schematic, relevance, true),
(schematic, conduciveness, true),
(schematic, urgency, true),
(schematic, control, true),
(suddenness, novelty, true),
(familiarity, novelty, true),
(novelty, relevance_detector, true),
(intrinsic_pleasantness, relevance_detector, true),
(relevance, relevance_detector, true),
(relevance_detector, conduciveness, guard1),

```
(relevance_detector, urgency, guard1),
(conduciveness, implication_detector, true),
(urgency, implication_detector, true),
(implication_detector, control, guard2),
(implication_detector, power, guard2),
(control, coping_potential_detector, guard3),
(power, coping_potential_detector, guard3)}
```

`L=ACTIVATION`, i.e., all dependencies represent that the first process *activates* the second.

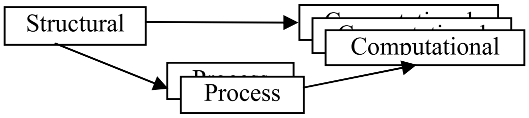

Figure 1. Three possible mappings between structural, process and computational models of emotion.

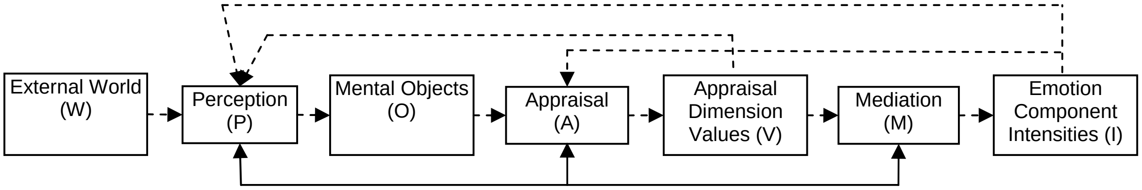

Figure 2. Graphical overview of the assumed structure of appraisal underlying our formalism. Dotted arrows denote potential inputs for processes, while normal arrows denote potential process dependencies. The external world contains events that can be perceived. Perception processes perceive event, appraisals 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.

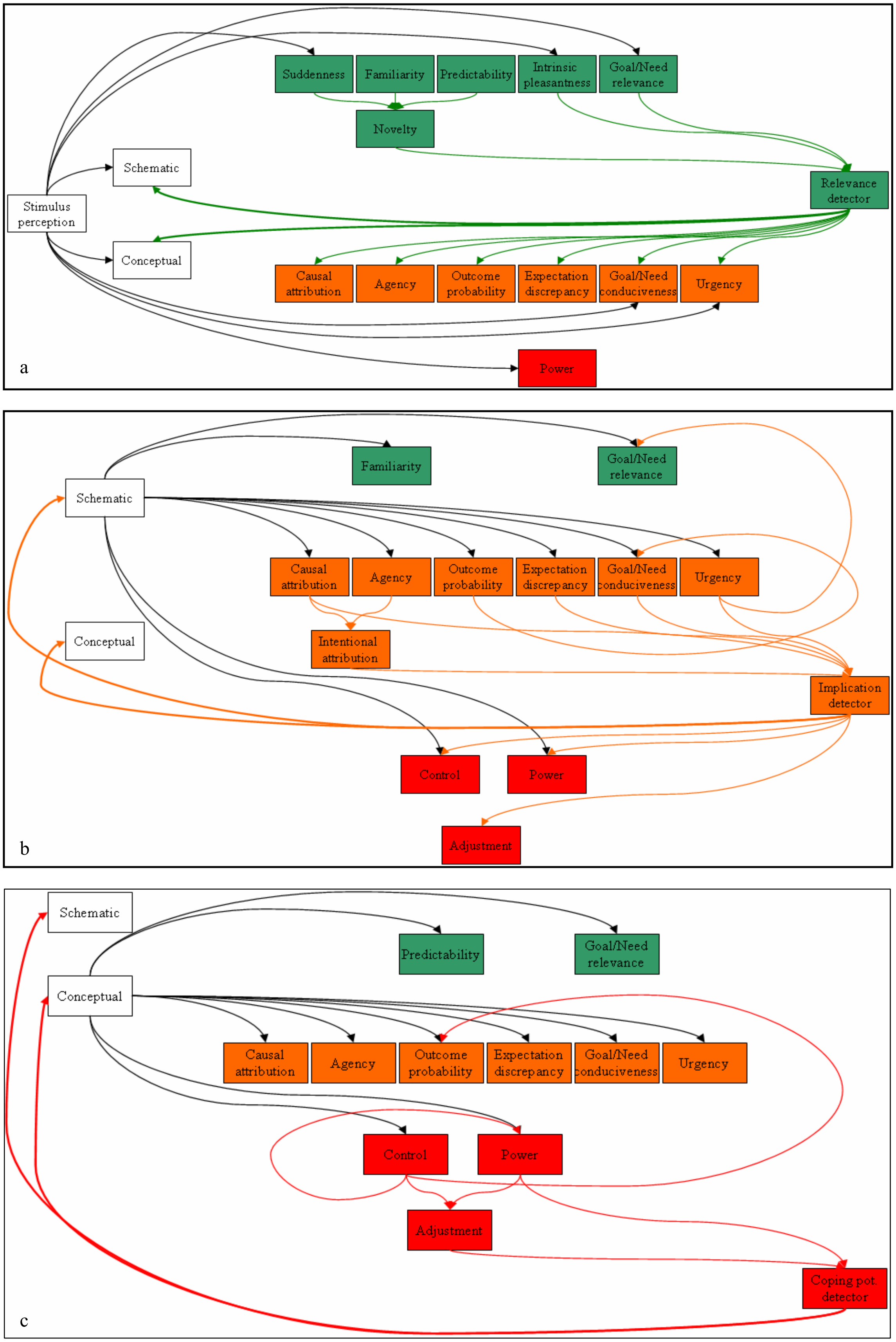
Suddenness
Familiarity
Predictability
Intrinsic pleasantness
Goal/Need relevance
Novelty
Schematic
Stimulus perception
Relevance detector
Conceptual
Causal attribution
Agency
Outcome probability
Expectation discrepancy
Goal/Need conduciveness
Urgency
Power
a
Familiarity
Goal/Need relevance
Schematic
Causal attribution
Agency
Outcome probability
Expectation discrepancy
Goal/Need conduciveness
Urgency
Conceptual
Intentional attribution
Implication detector
Control
Power
Adjustment
b
Schematic
Predictability
Goal/Need relevance
Conceptual
Causal attribution
Agency
Outcome probability
Expectation discrepancy
Goal/Need conduciveness
Urgency
Control
Power
Adjustment
Coping pot. detector
c

Figure 3a,b and c. 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).

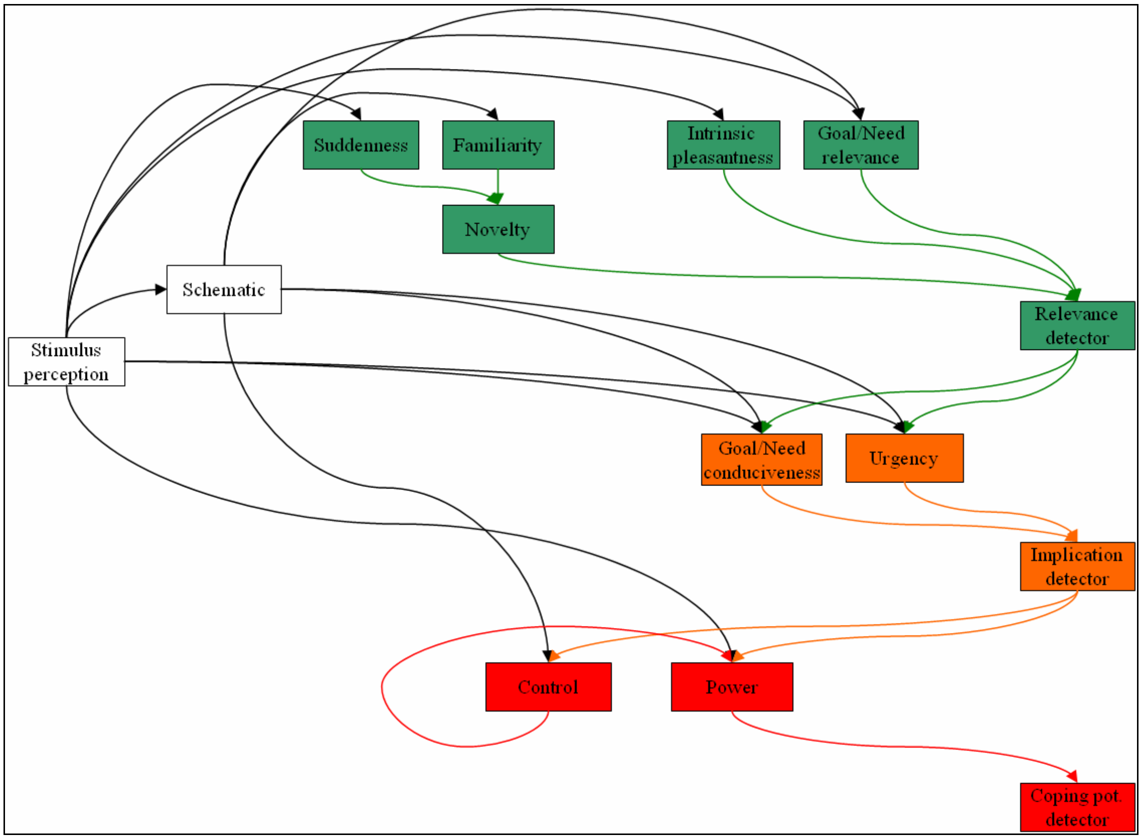

Figure 4. Graphical representation of the specification of PacMan's appraisal structure.

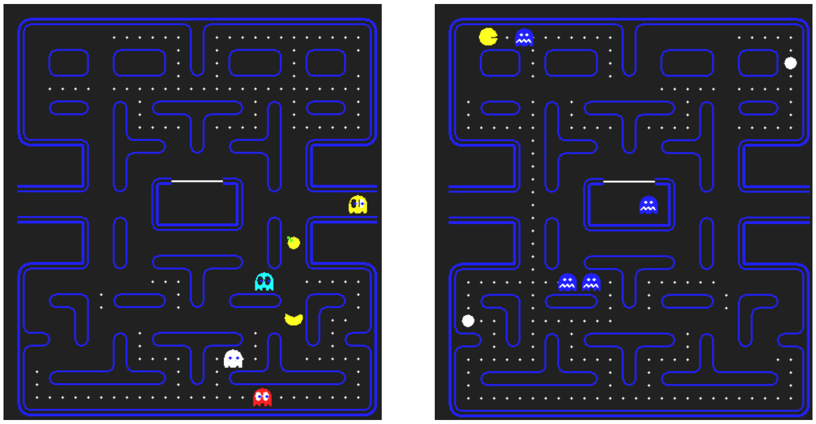

Figure 5. PacMan screen shots: chasing fruit (left), chasing an edible ghost (right).

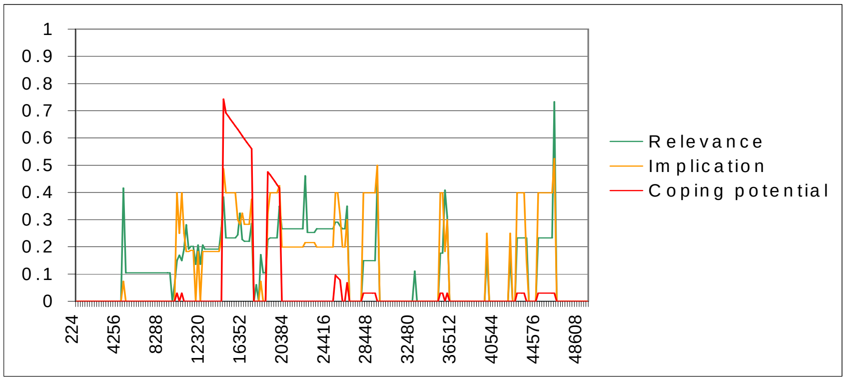

Figure 6a. PacMan using bi-polar variables. Time in milliseconds is on the x-axis. Activity of the levels of appraisal on the y-axis.

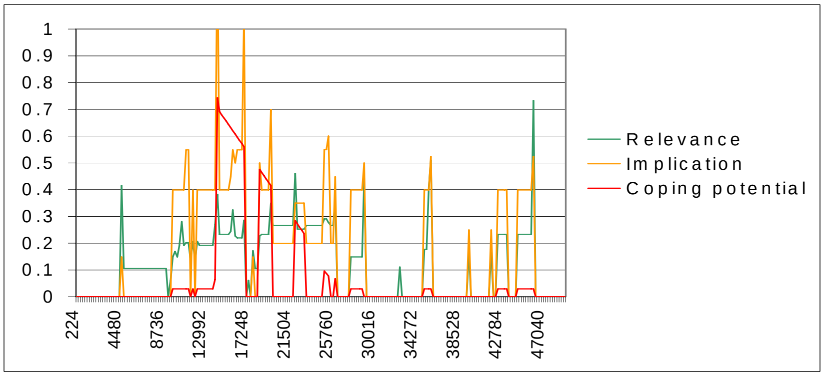

Figure 6b. PacMan without bi-polar variables.