To appear in *International Journal of Synthetic Emotions*, Vol. 2(1), 2011

GUIDELINES FOR DESIGNING COMPUTATIONAL MODELS OF EMOTIONS

Eva Hudlicka

Psychometrix Associates, Inc.
1805 Azalea Drive, Blacksburg, VA, US
hudlicka@ieee.org

Abstract

The past 15 years have witnessed a rapid growth in computational modeling of emotion and cognitive-affective architectures. Emotion models and architectures are being built both to elucidate the mechanisms of emotions, and to enhance believability and effectiveness of synthetic agents and robots. Yet in spite of the many emotion models developed to date, there is a lack of consistency, and clarity, regarding what exactly it means to 'model emotions'. More importantly, there are no systematic guidelines for the development of computational models of emotions. The purpose of this paper is to attempt to deconstruct the often vague term 'emotion modeling' by (1) suggesting that we view emotion models in terms of two fundamental categories of processes: emotion generation and emotion effects; and (2) identifying some of the computational tasks necessary to implement these processes. The paper then discusses how these computational building blocks can provide a basis for the development of more systematic guidelines for affective model development. The paper concludes with a description of an affective requirements analysis and design process for developing affective computational models in agent architectures. The aim of the proposed computational analytical framework is to contribute to a more systematic approach to the design of affective models, by specifying both the constituent computational building blocks, and the alternatives available for their implementation.

Keywords: Emotion, Modeling, Computational Affective Models, Core Affective Processes, Cognitive-Affective Architecture, Agent, Design Guidelines

Table of Contents

List of Figures

List of Tables

1.0 Introduction and Objectives

The past 15 years have witnessed a rapid growth in computational models of emotion and affective agent architectures. Researchers in cognitive science, AI, HCI, robotics, and gaming are developing 'models of emotion' for theoretical research regarding the nature of emotion, as well as a range of applied purposes: to create more believable and effective synthetic characters and robots, and to enhance human-computer interaction.

Yet in spite of the many stand-alone emotion models, and the numerous affective agent and robot architectures developed to date, there is a lack of consistency, and lack of clarity, regarding what exactly it means to 'model emotions' (Hudlicka, 2008b). 'Emotion modeling' can mean the dynamic generation of emotion via black-box models that map specific stimuli onto associated emotions. It can mean generating facial expressions, gestures, or movements depicting specific emotions in synthetic agents or robots. It can mean modeling the effects of emotions on decision-making and behavior selection. It can also mean including information about the user's emotions in a user model in tutoring and decision-aiding systems, and in games.

There is also a lack of clarity regarding what affective states are modeled. The term 'emotion' in affective models can refer to emotions proper (short, transient states), moods, mixed states such as attitudes, and frequently states that are not considered by psychologists as emotions (e.g., confusion, flow).

Emotion models also vary greatly regarding exactly which of the many roles ascribed to emotions are modeled. These include goal management and goal selection, resource allocation and subsystem coordination, and communication and coordination among agents, and among virtual agents and humans.

One of the consequences of this terminological vagueness is that when we begin to read a paper addressing 'emotion modeling', we don't really know what to expect. The paper could just as easily describe details of facial expression generation, affective speech synthesis, black-box models mapping domain-specific stimuli onto emotions, or decision-utility formalisms evaluating behavioral alternatives. A more serious consequence is a lack of design guidelines regarding how to model a particular affective phenomenon of interest: What are the computational tasks that must be implemented? Which theories are most appropriate for a given model? What are the associated representational and reasoning requirements, and alternatives? What data are required from the empirical literature?

The lack of consistent, clear terminology also makes it difficult to compare approaches, in terms of their theoretical grounding, their modeling requirements, and their theoretical explanatory capabilities and their effectiveness in particular applications.

The purpose of this paper is to attempt to deconstruct the vague term 'emotion modeling' by: (1) suggesting that we view emotion models in terms of two fundamental categories of processes: *emotion generation* and *emotion effects;* and (2) identifying some of the fundamental computational tasks necessary to implement these processes. These 'model building blocks' can then provide a basis for the development of more systematic guidelines for emotion modeling, theoretical and data requirements, and representational and reasoning requirements and alternatives. Identification of a set of generic computational tasks also represents a good starting point for a more systematic comparison of alternative approaches and their effectiveness. A systematic identification of the required building blocks also helps answer more fundamental questions about emotions: What are emotions? What is the nature of their mechanisms? What roles should they play in synthetic agents and robots? These computational building blocks can thus begin to serve as basis for what Sloman calls "architecture based definition of emotion" (Sloman, Chrisley, & Scheutz, 2005).

A number of researchers have addressed the issue of systematizing emotion modeling, both at the individual task level, and at the architecture level. Reilly (2006) outlined some of the computational tasks addressed in this paper, and focused in particular on an analysis of the approaches to intensity calculation and combining similar emotions. Broekens and colleagues (Broekens, DeGroot, & Kosters, 2008) explored the use of an abstract set-theoretic formalism to systematically compare several existing theories of cognitive appraisal. Cañamero discussed design requirements for affective agents, focusing on the role of emotion in action selection (L. D. Cañamero, 2001).

Lisseti and Gmytrasciewicz identified a number of high-level components of emotion required for computational models, in their Affective Knowledge Representation scheme (Lisetti & Gmytrasiewicz, 2002). In terms of architectures, Sloman and colleagues have done extensive work in exploring the architectural requirements for different classes of adaptive behavior, including requirements for different types of emotions, within the context of their CogAff architecture (Sloman et al., 2005). Ortony and colleagues have proposed a high-level design for an architecture that explicitly models emotion, and also includes a brief discussion of affective states and traits as parameters influencing processing (Ortony, Norman, & Revelle, 2005). Fellous (2004) addressed the need and requirements for emotions in synthetic agents from a neuroscience perspective.

The analytical framework presented in this paper (consisting of the core affective processes and the associated computational tasks necessary to implement them), and the set of guidelines defined for designing affective models, follows in this tradition. However, by providing a concrete set of the computational tasks required, and alternatives for their implementation, this paper aims to go beyond the existing work in affective modeling, and provide the first steps towards the development of more systematic guidelines for the design of computational models of emotion.

The paper is organized as follows. Section 2 provides an overview of emotion research in psychology. Section 3 discusses the theoretical foundations available for computational models of emotions. Section 4 introduces the computational perspective on emotion modeling, discussing the core affective processes (emotion generation and emotion effects modeling), the distinct domains that must be defined to implement these processes and the computational tasks required, and briefly addresses the associated representational and reasoning requirements. Section 5 then describes a systematic process for affective requirements analysis and model design. Section 6 provides a summary and conclusions.

2.0 Background on Emotion Research in Psychology

Regardless of the purpose of a particular affective computing research or development project, it is important for the researcher or practitioner to be familiar with the emotion research literature in psychology. This extensive body of research provides definitions and terminology, theories and conceptual models, and vast quantities of experimental data. Familiarity with these theories, models and data is critical for the development of affective computational models and affective user models, as well as for emotion recognition by machines, and generation of affective expressions in agents and robots. This is the case for applied models, whose aim is to enhance human-computer interaction or control the behavior of synthetic agents or robots. It is even more critical for research models, whose aim is to elucidate the nature of emotion, and the mechanisms of affective processes and phenomena. Knowledge of existing data and theories also provides a basis for evaluation and validation of computational models. Furthermore, familiarity with psychological emotion research facilitates the cross-disciplinary communication necessary for continued progress in computational emotion research, and emotion research in general. In this section we provide an overview of emotion research in psychology, as it relates to computational affective modeling.

2.1 Definitions: What are Emotions?

When searching for a definition of emotions, it is interesting to note that most definitions involve descriptions of characteristics (e.g., fast, undifferentiated processing) or roles, and functions (e.g., coordinating mechanisms for goal management in uncertain environments, communicative mechanisms for facilitating social interaction, hardwired responses to critical stimuli). The fact that we so often describe emotions in terms of their characteristics, rather than their essential nature, underscores our lack of understanding of these complex phenomena. Nevertheless, a number of emotion researchers in psychology do appear to agree on a high-level definition of emotions, and view emotions as *states that reflect evaluative judgments of the environment, the self and other social agents, in light of the organism's goals and beliefs, which motivate and coordinate adaptive behavior.* Note that the terms 'goals' and 'beliefs' are used in a generic sense: goals reflecting desirable states, and beliefs reflecting current knowledge. The term 'goal' in this discussion therefore covers any representation of desired states, and includes conscious and unconscious, explicit or implicit, and innate or learned goals.

2.2 Taxonomy of Affective States and Traits

The term ‘emotion’ itself is problematic. On the one hand, it depicts emotions in a generic, folk-psychology sense we all presume to understand, and which subsumes many types of affective states. On the other hand, it has a specific meaning in the emotion research literature, referring to transient states, lasting for seconds or minutes, typically associated with well-defined triggering cues and characteristic patterns of expressions and behavior. (More so for the simpler, fundamental emotions than for complex emotions with strong cognitive components.)

Emotions are often further categorized into *basic* emotions (P. Ekman, 1992; Izard, 1977; Panskepp, 1998; Plutchik, 1984)) (also referred to as fundamental) and, for lack of a better word, *non-basic or complex* emotions. The inelegant terms non-basic and complex reflect the difficulty of classifying these emotions under a single, descriptive term. The set of non-basic emotions includes the important social (also termed self-conscious) emotions, such as pride, shame, and guilt, but also all other complex emotions, including love, empathy, shadenfreude, humiliation, contempt, and many others.

This categorization reflects differences in the degree of cognitive complexity associated with specific emotions, the universality of the triggering stimuli and the expressive and behavioral manifestations, and the degree to which an explicit representation of the agent’s self within its social milieu is required. While not universally accepted in emotion research (Ortony & Turner, 1990) and (P. Ekman & Davidson, 1994), this categorization, is nevertheless useful for affective modeling, and is adopted in the discussion below.

Emotions can be contrasted with other terms describing affective phenomena: *moods*, sharing many features with emotions but lasting longer (hours to months), and having less differentiated triggers and manifestations; *affect*, an undifferentiated positive or negative state associated with generic behavior tendencies (approach, avoid); and *feelings,* a problematic and ill-defined construct from a modeling perspective. (Averill points out that “feelings are neither necessary nor sufficient conditions for being in an emotional state” (Averill, 1994)). Emotions, moods, affect and feelings all refer to affective states, i.e., transient affective activity. Some models also represent stable affective personality *traits* (e.g., extraversion, neuroticism), or a variety of 'mixed' mental states that involve both cognitive and affective components (e.g., attitudes). Figure 1 illustrates the taxonomy of affective states and traits described above.

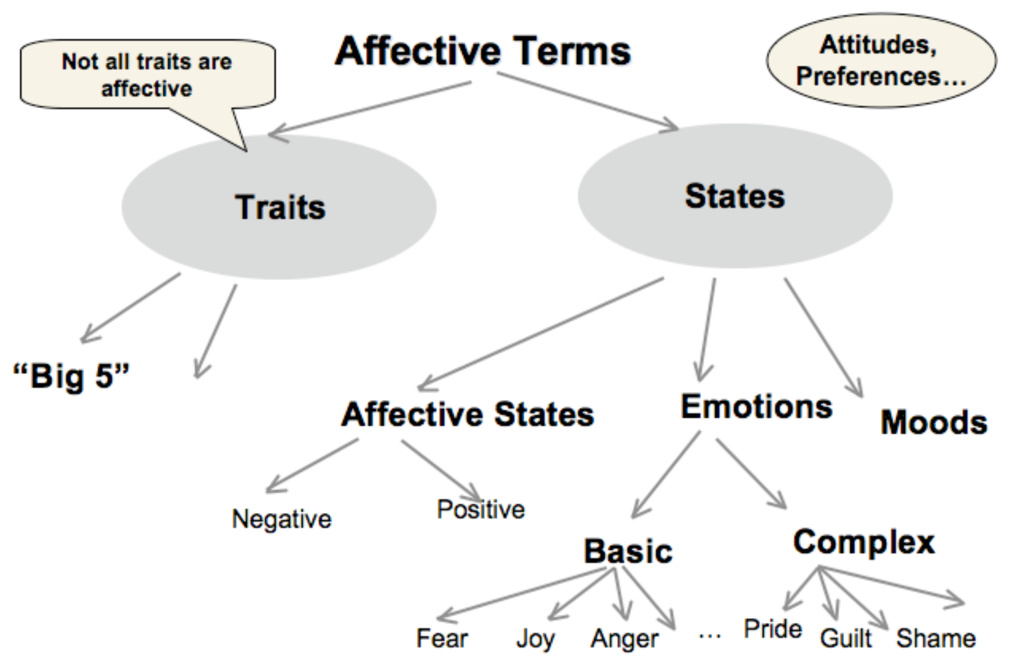


Figure 1: Taxonomy of Affective States and Traits

2.3 Multi-Modal Nature of Emotions

A key aspect of emotions is their multi-modal nature. Emotions in biological agents are manifested across four distinct, but interacting, modalities. The most familiar is the *behavioral / expressive* modality, where the expressive and action-oriented characteristics are manifested; e.g., facial expressions, speech, gestures, posture, and behavioral choices. Closely related is the *somatic / physiological modality* - the neurophysiological substrate making behavior (and cognition) possible (e.g., heart rate, neuroendocrine effects, blood pressure). The *cognitive / interpretive* modality is most directly associated with the evaluation-based definition provided above, and emphasized in the current cognitive appraisal theories of emotion generation, discussed below. The most problematic modality, from a modeling perspective, is the *experiential / subjective* modality: the conscious, and inherently idiosyncratic, experience of emotions within the individual.

While the current emphasis in emotion modeling is on the cognitive modality (involved in appraisal) and the behavioral /expressive modality (manifesting emotions in agents), it is important to recognize that both the physiological, and the experiential modalities, also play critical roles (Izard, 1993). Table 1 shows examples of the multi-modal signatures of two of the 'basic' emotions: anger and fear.

Table 1: Examples of Multimodal Signatures of Emotions

	Anger
Trigger(s)	Progress toward goals hindered, especially by another agent
Cognitive Biases	Strong effect on attention capture and focus Attribution of hostility to other agents Overestimates of own chances of success Tendency to try alternative strategies Reduced perception of risk & increased risk tolerance
Physiological Correlates	Overall objective: mobilize and sustain high energy levels required for possible aggression
ANS Correlates	Larger heart rate acceleration (than happiness); Larger increase in finger temperature (than fear); Higher diastolic blood pressure (than fear); Greater peripheral resistance (than fear); Larger increase in heart rate (than disgust) (Levenson, 1992)
Facial Expr.	Lowering brows; raising upper eyelids; tightening & pressing together of lips
Speech Characteristics	Fast rate; High voice intensity; high pitch; wide range of pitch; abrupt pitch changes
Behavioral Manifestations	Eagerness to act Tendency towards aggression

	Fear
Trigger(s)	Perceived danger to important self- or other-protective goals
Cognitive Biases	Tunnel vision (attentional narrowing focusing on the source of the threat) Threat focus in attention and interpretation Faster detection of threatening stimuli Interpretation of ambiguous stimuli as threatening
Physiological Correlates	Mobilization of energy level to prepare for action; energy 'spike'
ANS Correlates	Larger heart rate acceleration (than happiness); Larger skin conductance increase (than happiness); Lower diastolic BP (than anger); Decreased finger temperature (Levenson, 1992)
Facial Expr.	Raised eyebrows and eyelids (eyes wide open); lips stretched and pressed together
Speech	Faster rate; higher pitch; wider pitch range; medium intensity
Goals	Self protection; Protection in general
Behavioral Manifestations	Flight and avoidance; Protective behavior; Low intensities (inhibitory behavior, freezing), high intensities (fleeing) (Panskepp, 1998, p. 208)

2.4 What are the Roles of Emotions?

Emotions play a number of important roles in biological agents, by helping to regulate homeostasis, reproductive and survival behaviors, and by enabling adaptive behavior in complex and uncertain environments in general, including social behavior. Emotion roles can be categorized into two broad groups: *intrapsychic and interpersonal* (see Figure 2). Both are discussed in more detail below. Table 2 provides a summary of both the intrapsychic and the interpersonal roles of emotions.

The currently accepted view regarding the evolution and utility of emotions is that their primary role is to ensure survival, by rapidly detecting survival-critical stimuli in the environment, and by preparing the organism for a coordinated execution of the necessary behavioral responses (P. Ekman & Davidson, 1994; N. Frijda, 2008; Plutchik, 1984). We know however that emotions can also become highly maladaptive, even dangerous, both to the individual experiencing or manifesting the emotions, and to other agents in its social environment. Until relatively recently, the maladaptive and pathological nature of emotions received the most attention in psychology. However, the focus in affective modeling is primarily on the adaptive roles of emotions, although some models have been developed that attempt to identify the mechanisms of maladaptive affective processing; e.g., the mechanisms underlying anxiety disorders ((Hudlicka, 2008a)).

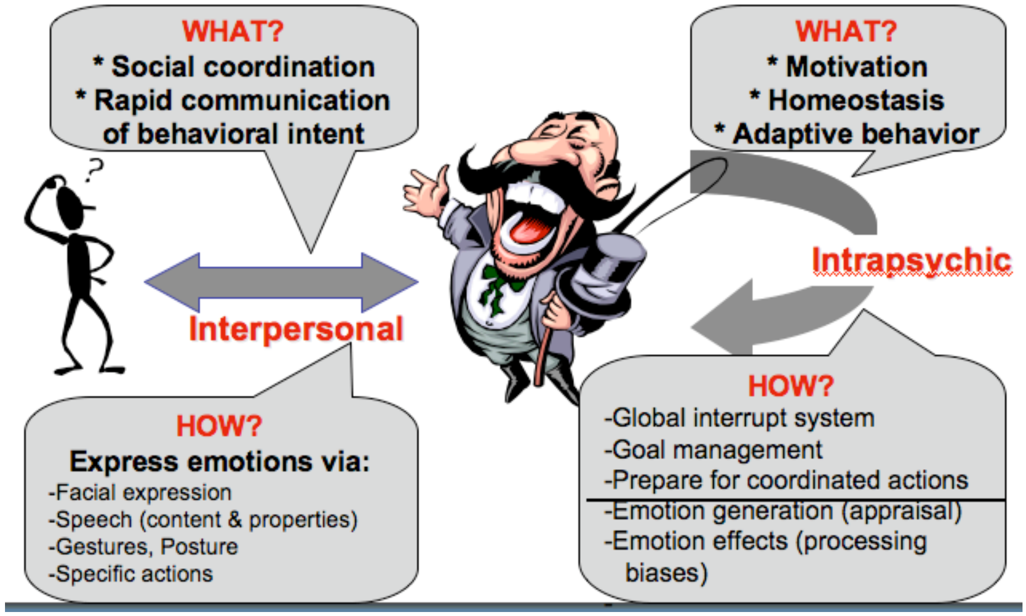


Figure 2: Two Categories of Emotion Roles: Interpersonal (Social) and Intrapsychic

Table 2: Summary of Intrapsychic and Interpersonal Roles of Emotions

Intrapsychic Roles
• Rapid detection & processing of salient stimuli (e.g., avoid danger, get food)
• Triggering, preparation for, and execution of fixed behavioral patterns necessary for survival (e.g., fight, freeze, flee)
• Rapid resource (re)allocation & mobilization
• Coordination of multiple systems (perceptual, cognitive, physiological)
• Implementation of systemic biases biasing processing in a particular direction (e.g., threat detection, self-focus)
• Interruption of on-going activity & (re)prioritization of goals
• Motivation of behavior via reward & punishment mechanisms
• Motivation of learning via boredom & curiosity
Interpersonal Roles
• Communication of internal state via non-verbal expression and behavioral tendencies (e.g., frown vs. smile, inviting vs. threatening gestures & posture)
• Communication of status information in a social group (dominance & submissiveness)
• Mediation of attachment
• Communicate acknowledgment of wrong-doing (guilt, shame) in an effort to repair relationships and reduce possibility of aggression

3.0 Theoretical Foundations for Computational Models of Emotions

A number of theories have been developed in psychology that have direct relevance for computational affective modeling. Below we discuss both the broad theoretical perspectives on emotions, and specific theories relevant to modeling emotion generation and emotion effects, with an emphasis on the cognitive modality.

Adherence to a particular theoretical perspective, and the use of specific theories, help ensure that the models are developed in a theoretically and empirically grounded manner. This then facilitates identification of the appropriate empirical data, model validation, communication with other researchers, and potential sharing of model components. In addition, adherence to particular theoretical perspectives helps advance our understanding of affective processes, by enabling computation-based elaborations of high-level theories, which often reveal gaps or contradictions, thus enabling the development of more accurate theories of affective phenomena.

3.1 Theoretical Perspectives

Emotions represent complex, and often poorly understood, phenomena. It is therefore not surprising that a number of distinct theories have evolved over time, to explain a specific subset of these phenomena, or to account for a particular subset of the observed data. Three of the most established theoretical perspectives, and those most relevant for computational affective modeling, are described below.

Discrete theories of emotions emphasize a small set of discrete or fundamental emotions. The underlying assumption of this approach is that these fundamental, discrete emotions are mediated by associated neural circuitry, with a large innate, 'hardwired' component. Different emotions are then characterized by stable patterns of triggers, behavioral expression, and associated distinct subjective experiences. The emotions addressed by these theories are typically the 'basic' emotions; joy, sadness, fear, anger, and disgust. Because of its emphasis on discrete categories of states, this approach is also termed the *categorical approach* (Panskepp, 1998). For modeling purposes, the semantic primitives representing emotions in affective models are the basic emotions themselves.

An alternative method of characterizing affective states is in terms of a small set of underlying dimensions, that define a space within which distinct emotions can be located. This *dimensional perspective* describes emotions in terms of two- or three-dimensions. The most frequent dimensional characterization of emotions uses two dimensions: valence and arousal (Russell, 2003; Russell & Barrett, 1999; Russell & Mehrabian, 1977). Valence reflects a positive or negative evaluation, and the associated felt state of pleasure (vs. displeasure), as outlined in the context of undifferentiated affect above. Arousal reflects a general degree of intensity or activation of the organism. The degree of arousal reflects a general readiness to act: low arousal is associated with less energy, high arousal with more energy. Since this 2-dimensional space cannot easily differentiate among emotions that share the same values of arousal and valence, e.g., anger and fear, both characterized by high arousal and negative valence, a third dimension is often added. This is variously termed dominance or stance. The resulting 3-dimensional space is often referred to as the PAD space (Mehrabian & 1995) (pleasure (synonymous with valence), arousal, dominance). The representational semantic primitives within this theoretical perspective are thus these 2 or 3 dimensions.

The third view emphasizes the distinct components of emotions, and is often termed the componential view (Leventhal & Scherer, 1987). The 'components' referred to in this view are both the distinct modalities of emotions (e.g., cognitive, physiological, behavioral, subjective) and also the components of the cognitive appraisal process. These are referred to as appraisal dimensions or appraisal variables, and include novelty, valence, goal relevance, goal congruence, and coping abilities. A stimulus, whether real or imagined, is analyzed in terms of its meaning and consequences for the agent, to determine the affective reaction. The analysis involves assigning specific values to the appraisal variables. Once the appraisal variable values are determined by the organism's evaluative processes, the resulting vector is mapped onto a particular emotion, within the n-dimensional space defined by the n appraisal variables. The semantic primitives for representing emotions within this model are thus these individual appraisal variables.

It must be emphasized that these theoretical perspectives should not be viewed as competing for a single ground truth, but rather as distinct perspectives, each arising from a particular research tradition (e.g., biological vs. social psychology), focusing on different sets of affective phenomena, considering distinct levels of resolution and fundamental components (e.g., emotions vs. appraisal variables as the distinct primitives), and using different experimental methods (e.g., factor analysis of self-report data vs. neuroanatomical evidence for distinct processing pathways). The different perspectives also provide different degrees of support for the distinct processes of emotion; e.g., the componential theories provide extensive details about cognitive appraisal.

Until such time as emotions are fully understood and explained, it is best to view the three theoretical perspectives outlined above as alternative explanations, each with its own set of explanatory powers and scope, and supporting data, analogously, perhaps, to the wave vs. particle theory of light, as suggested by Picard (Picard, 1997). Table 3 provides a summary of these three theoretical perspectives.

Table 3: Summary of the Three Dominant Theoretical Perspectives Available for Affective Modeling

(Note that the OCC theory fits best within the componential view, because the abstract features comprising the emotion eliciting conditions are analogous to the appraisal variables of the componential theories.)

	Discrete / Categorical	**Dimensional**	**Componential**
Semantic Primitives	**Basic emotions** Small number of 'hardwired' emotions exists (basic emotions), with characteristic neuromotor patterns that prepare for specific, species-relevant behaviors	**2 or 3 dimensions** (valence and arousal, or valence, arousal, dominance) (often the term pleasure is used instead of valence, thus PAD)	**Appraisal variables** Emotions result from the (parallel) evaluation of a number of *appraisal variables* (e.g., novelty, valence, goal congruence, agency). Appraisals of individual variables represent components of emotions.
Semantic Primitives Correspond to	Basic emotions	Underlying dimensions of felt mood (or core affect)	Interpretive features of the stimulus, and the stimulus-agent relationship
Number / Type of Emotions Possible	Small set of basic emotions (e.g., joy, sadness, anger, fear, disgust, surprise). Complex/social emotions thought to be combinations of basic emotions.	A large set of emotions defined by combinations of different values of {valence & arousal} or {valence, arousal, & dominance}. Larger set than basic dimensions but not as large as the set defined by the appraisal variables of the component process model.	Very large set of emotions defined by the possible values of all of the appraisal variables
Degree of Elaboration of Appraisal Process	Low	Low (dimensions originally derived from mood, thus with a strong physiological component, vs. cognitive component)	Very high
Degree of Detail About Affective Dynamics	Low / Qualitative	Medium / Focus on arousal	Some / Qualitative
Representative Theorists	Tomkins, Ekman, Izard, Panskepp	Wundt, Mehrabian, Russell, Lang, Watson & Tellegen	Frijda, Scherer, Smith, Ellsworth, Kirby, Roseman, Ortony et al. (OCC)

3.2 Theories of Emotion Generation and Emotion Effects

Below we provide an overview of the dominant theories available for modeling emotion generation and emotion effects.

3.2.1 Emotion Generation

Emotion generation is an evolving, dynamic process that occurs across the multiple modalities, discussed above, with complex feedback and interactions among them. While all modalities are involved, our understanding of these phenomena is best within the cognitive modality, and most existing models of emotion generation implement *cognitive appraisal* (exceptions do exist, e.g., (Breazeal & Brooks, 2005; L. Cañamero, 1997; L. Cañamero & Avila-Gracia, 2007; Velásquez, 1999). The discussion below is therefore limited to cognitive appraisal, recognizing that the current cognitive bias may well be an example of "looking for the key under the lamp because there is light there". In addition, as embodied agents become more complex, there may be increasing need to include non-cognitive modalities in emotion generation.

All cognitive appraisal theorists emphasize the critical role that cognition plays in generating the subjective emotional experience, by mediating the interpretations required for the evaluative judgments involved in generating emotion. Figure 3 illustrates the appraisal process, indicating how the same stimulus can produce distinct emotions, depending on the differences in the individual's interpretations. Appraisal theories have their roots in antiquity and have gone through a number of iterations since then. Many researchers over the past four decades have contributed to the current versions of cognitive appraisal theories (Arnold, 1960; N. H. Frijda, 1986; Lazarus, 1984; Mandler, 1984; I.J. Roseman & Smith, 2001; K. Scherer, Schorr, & Johnstone, 2001; Smith & Kirby, 2001).

Appraisal theorists recognize that appraisal processes vary in complexity and cognitive involvement, from: low-level, 'hardwired', to complex, culture-specific and idiosyncratic triggers. Three interconnected levels are typically proposed: sensorimotor, schematic, and conceptual. (Similar tri-level organization has also been proposed for cognitive-affective architectures in general (Ortony et al., 2005; Sloman et al., 2005)).

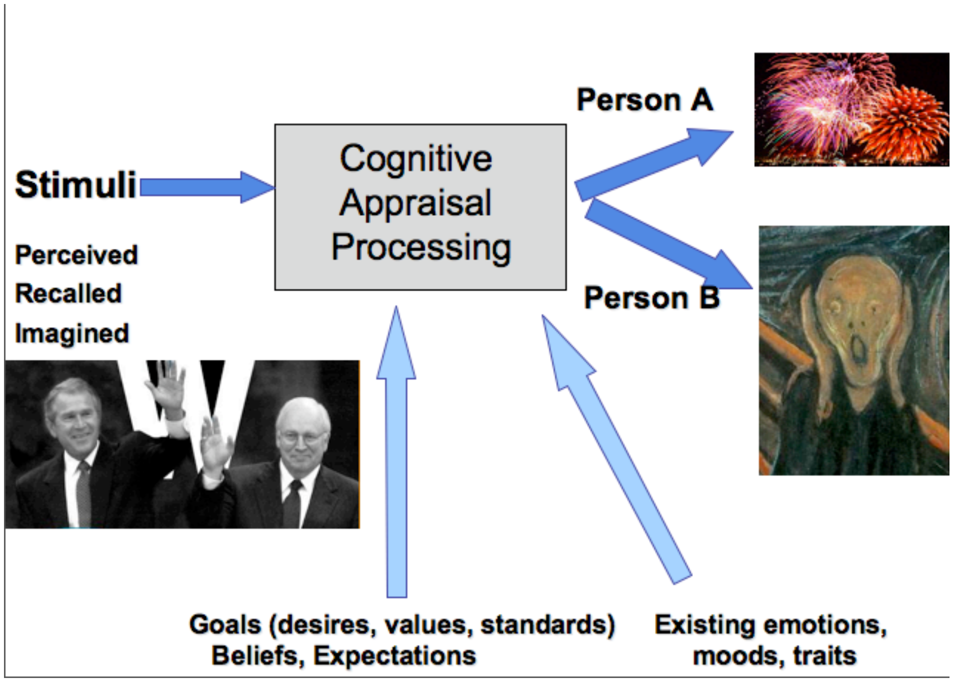


Figure 3: High-Level Illustration of Emotion Generation via Cognitive Appraisal

The most influential appraisal theories in computational modeling are those that are cast in 'computation-friendly' terms. The first of these were the theories proposed by Ortony and colleagues, now referred to as the OCC theory (Ortony, Clore, & Collins, 1988) and Frijda (1986). The OCC theory is the most frequently implemented theory of cognitive appraisal in computational models of emotion (Andre, Klesen, Gebhard, Allen, & Rist, 2000; Aylett, 2004; Bates, Loyall, & Reilly, 1992; de Rosis, Pelachaud, Poggi, Carofiglio, & De Carolis, 2003; El-Nasr, Yen, & Ioerger, 2000; Elliot, 1992; Gratch & Marsella, 2004; Loyall, 1997; Prendinger, Saeyor, & Ishizuka, 2004; W. S. R. Reilly, 1996; Staller & Petta, 1998). Several models have used Frijda's appraisal theory (N. H. Frijda, 1986; N. H. Frijda & Swagerman, 1987). More recently, appraisal theories of Scherer, Roseman, and Smith and Kirby, have begun to be used as the basis of computational models (I.J. Roseman, 2001; K. Scherer et al., 2001; Smith & Kirby, 2000).

There is a considerable degree of overlap among these theories, in terms of the features used to evaluate the emotion eliciting triggers (e.g., desirability, likelihood, responsible agent), but also some significant differences in the emotion taxonomies they provide, and the sets of domain-independent features or variables used to characterize the triggering stimuli. These similarities and differences will be discussed below. The remainder of this section discusses the OCC theory, and the componential appraisal theories, in more detail

3.2.1.1 OCC Theory of Cognitive Appraisal

The theory of cognitive appraisal developed by Ortony and colleagues (Ortony et al., 1988) (OCC) remains the most frequently implemented theory in computational models of emotion generation (e.g., (Andre et al., 2000; Bates et al., 1992; W. S. N. Reilly, 2006). It provides a rich taxonomy of triggers and the resulting emotions, a distinguishing feature of the OCC theory, which emphasizes fundamental distinctions among three types of triggers, and corresponding types of emotions (refer to

Figure 4). An OCC-based appraisal process proceeds through a sequence of steps as it classifies a trigger within this taxonomy, eventually generating a specific emotion (refer to lower part of

Figure 4).

Within the OCC theory, emotions reflect an agent's *valenced reactions* to three different types of triggers: *events, actions by other agents,* and *objects*. As with other theories, emotions are triggered when the agent's goals are either achieved or hindered, or when the potential for achievement or hindrance exists. In the OCC theory however, the term goal has a specific meaning. The OCC model speaks broadly about concerns, but then divides these into three distinct categories of concerns, each associated with the different categories of triggers and emotions: (1) goals, (2) standards and norms for behavior, and (3) preferences and tastes. Each of these categories of concerns is associated with distinct triggers (events with goals; acts by agents with standards and norms; and object attributes with tastes and preferences), and produces different categories of emotions.

Event-based emotions reflect the agent's general well-being, and involve an evaluation of a triggering event with respect to the agent's goals. If an event is conducive to the agent's goals, it triggers a positive emotion. If it hinders a goal, it triggers a negative emotion. These emotions can be further differentiated, depending on whether the event concerns the self or another, whether it is occurring in the current time frame or anticipated in the future, and whether the event was anticipated or not. Emotions associated with the agent's own *well-being* include joy and distress. Emotions associated with what OCC call *fortunes of others* include happy-for, sorry-for, gloating and resentment. Emotions associated with future desirable or undesirable events are referred to as *prospect-based emotions,* and include hope and fear. Emotions associated with confirmation or disconfirmation of anticipated events include satisfaction, relief, fears-confirmed, and disappointment.

To further support the differentiated representation of emotions, OCC model offers a taxonomy of goals, based on earlier work of Schank and Abelson (Schank & Abelson, 1977). *Active pursuit goals* (A goals) reflect goals that can theoretically be achieved by the agent, and for which the agent has some plans, and a means of implementing those plans; e.g., write a grant proposal, wash dishes, go on vacation. *Interest goals* (I goals) are goals that the agent cannot directly achieve, but which represent states of the self, world or others that the agent wishes were true; e.g., have favorite actor win the Oscars. Finally, the *replenishment goals* (R goals) are goals that are achievable by the agent, but reflect on-going needs that are cyclical in nature, and typically reflect some homeostatic state of a particular variable; e.g., maintain adequate level of satiation, hydration, rest. These goals are often linked to basic drives in biological agents.

Acts by agents trigger agent-based *attribution emotions* reflecting approval or disapproval. These can reflect both the self, and other agents. Thus pride and shame reflect a positive / negative evaluation of own behavior, whereas admiration and reproach reflect a positive / negative evaluation of other agents' behavior. Generation of these emotions is mediated by behavioral standards, against which the agent's own behavior, or other agents' behavior, is compared.

Characteristics of objects trigger object-based *attraction emotions* such as like, and dislike. These are mediated by general attitudes used to appraise the object. If an object is appealing it triggers positive emotions, such as liking, if not, it triggers negative emotions, such as disliking or disgust.

Some emotions result from combinations of triggers or evaluation criteria. These are called *compound emotions.* For example, combining *well-being* and *attribution emotions* can produce: gratitude (admiration and joy), anger (reproach and distress), gratification (pride and joy), or remorse (shame and distress); combining attraction and attribution emotions can produce: love (admiration and liking), or hate (reproach and disliking) (Elliot, 1992; O'Rorke & Ortony, 1994).

Different values of the evaluation criteria thus define specific emotions. Table 4 lists the emotions defined within the OCC theory, in terms of the values of the domain-independent evaluation criteria. The bottom part of

Figure 4 illustrates how the compound emotion ‘anger' would be derived, within the OCC taxonomy of emotions.

In addition to defining emotions in terms of their detailed evaluative structure, their *cognitive structure*, the OCC theory also provides a basis for emotion intensity calculation. Intensity is calculated from the values of the *local variables* (the evaluation criteria discussed above), whose specific values define distinct emotions, as well as *global* variables, which apply to all emotions. Global variables consist of the following: consist of the following: a sense of reality, proximity, unexpectedness and arousal. Global variables thus reflect both the physiological state of the agent (e.g., arousal) and properties of the triggers (e.g., proximity (both temporal and physical), unexpectedness, sense of reality).

The OCC model assumes the existence of context-specific thresholds for each emotion. When triggers appear that could potentially result in a particular emotion, that emotion will only be triggered if its intensity value exceeds the associated threshold. The degree to which this threshold is exceeded also influences the emotion intensity.

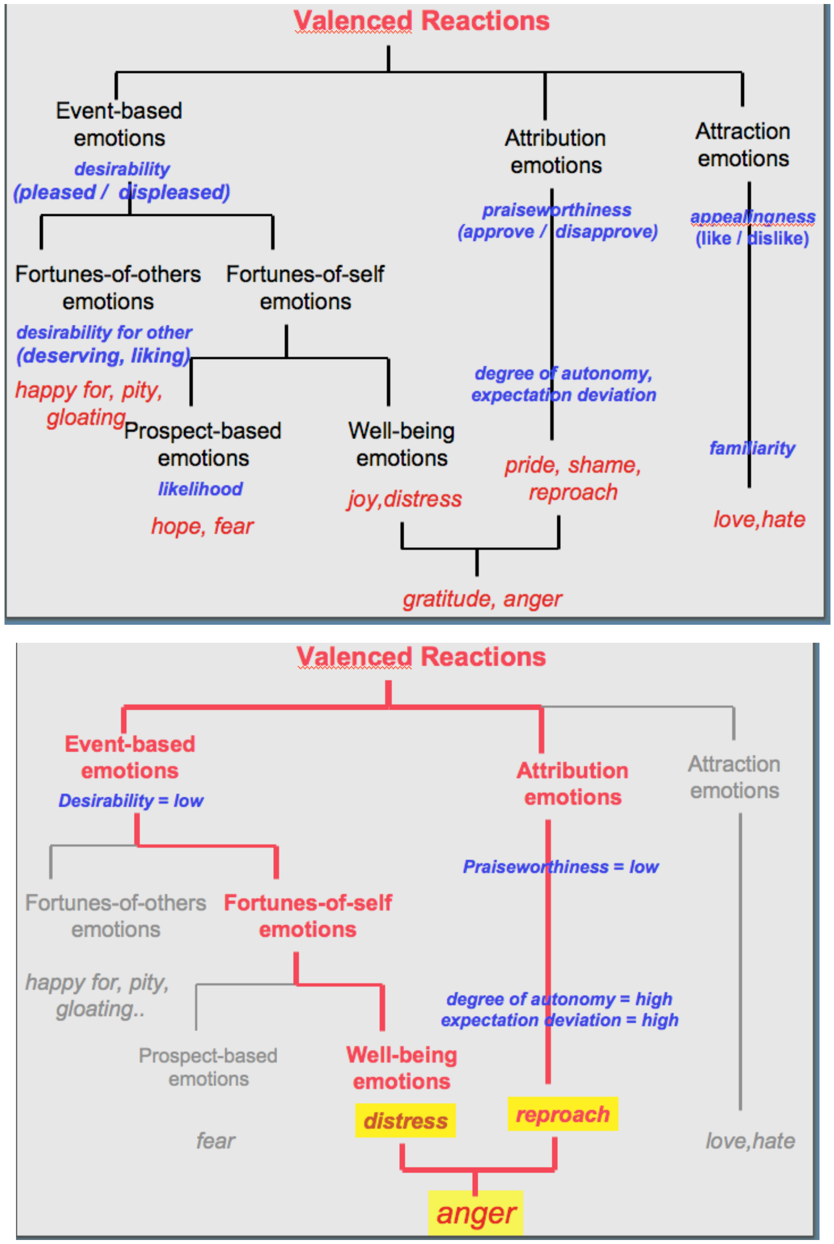


Figure 4: OCC Taxonomy of Emotion Triggers and Emotions

The top part of the figure shows the different categories of emotions and the types of evaluations necessary to derive them. The bottom part shows the derivation of anger, highlighting the activation of multiple branches in the taxonomy required.

Table 4: Definitions of Emotions in Terms of the OCC Triggers, Internal References and Evaluation Criteria (Local Variables)

(based on table 2.1 in Elliot, 1992 and (O'Rourke & Ortony, 1994)

Emotion	OCC Emotion Type	Trigger Type	Appraised w/ Respect	Evaluation criteria (Local Variables)
Simple Emotions (Evaluated with respect to single category of criteria)				
Joy	Well-being	Event affecting self	Goals	Desirability of event wrt goal
Distress	Well-being	Event affecting self	Goals	Undesirability of event wrt goal
Happy-for	Fortunes of others	Event affecting another agent	Goals	Pleased about a desirable event for another agent
Sorry-for	Fortunes of others	Event affecting another agent	Goals	Distressed about an undesirable event for another
Gloating	Fortunes of others	Event affecting another agent	Goals	Pleased about an event undesirable for another
Resentment	Fortunes of others	Event affecting another agent	Goals	Displeased about an event desirable for another
Hope	Prospect-based	Prospective Event	Goals	Pleased about a potential good event in the future
Fear	Prospect-based	Prospective Event	Goals	Distressed about a potential bad event in the future
Satisfaction	Confirmation	Prospective Event	Goals	Pleased because an expected good event occurred
Fears confirmed	Confirmation	Prospective Event	Goals	Distressed because an expected bad event occurred
Relief	Confirmation	Prospective Event	Goals	Pleased because an expected bad thing didn't happen
Disappointment	Confirmation	Prospective Event	Goals	Distressed because an expected bad thing did happen
Pride	Attribution	Act by self	Norms	Approving of own behavior
Shame	Attribution	Act by self	Norms	Disapproving of own behavior
Admiration	Attribution	Act by another	Norms	Approving of another's behavior
Reproach	Attribution	Act by another	Norms	Disapproving of another's behavior
Liking	Attraction	Entity qualities	Preferences	Finding the entity appealing
Disliking	Attraction	Entity qualities	Preferences	Finding the entity unappealing
Compound Emotions (Evaluated with respect to multiple categories of criteria)				
Gratitude	Well-being & Attribution	Event / Act by another	Goals / Norms	Joy + Admiration
Anger	Well-being & Attribution	Event / Act by another	Goals / Norms	Distress + Reproach
Gratification	Well-being & Attribution	Event / Act by self	Goals / Norms	Joy + Pride
Remorse	Well-being & Attribution	Event / Act by self	Goals / Norms	Distress + Shame
Love	Attribution & Attraction	Act by another / Quality of another	Norms / Preferences	Admiration + Liking
Hate	Attribution & Attraction	Act by another / Quality of another	Norms / Preferences	Reproach + Disliking

3.2.1.2 Componential Theories of Cognitive Appraisal

The componential theories of emotion emphasize the constituent components of emotions. There are several categories of "components": at the higher-level, components can refer to the different modalities. Within cognitive appraisal itself, components can also refer to the individual *appraisal variables*, which reflect evaluations along different evaluative criteria, analogously to some of the OCC theory's evaluative criteria discussed above. Like the OCC criteria, the appraisal variables represent domain-independent features of the triggers, or of the trigger-agent relationship. Appraisal variables include *novelty, valence, goal relevance and goal congruence, responsible agent, coping potential, and norms and values*. Specific configurations of the appraisal variable values then correspond to distinct emotions, and are often referred to as the agent's *appraisal state*. Within the componential appraisal theories, a given emotion is thus represented by a vector consisting of the values of the evaluated appraisal variables.

Several sets of such appraisal variables have been proposed by a number of cognitive appraisal theorists, including Scherer, Ellsworth, Frijda, Roseman, Smith, and Kirby. These sets offer very similar appraisal variables. Below we focus on the set proposed by Scherer, because it offers a larger set of variables, thereby allowing for a high-degree of differentiation of the distinct affective states (see (Ellsworth & Scherer, 2003; K.R Scherer, 2001). (Figure 5 shows the appraisal variables and Table 5 shows the relationships between appraisal variables values and different emotions).

Scherer divides the appraisal variables into four broad categories of evaluations, which occur in sequence, and correspond to evaluating the following criteria: *relevance, implications, coping potential, and norms* (refer to Figure 5). The first category (relevance) focuses largely on the properties of the stimulus. The later categories increasingly focus on the relationship between the triggering stimuli and the agent, and involve increasingly complex cognitive processing (although not necessarily conscious cognitive processing). This may require complex representations of the causal structure of the world, expected events, and the matching of internal schemas with incoming stimuli.

Once the appraisal variables are evaluated, the resulting vector corresponds to a point in the space defined by the dimensions corresponding to the individual appraisal variables, which corresponds to a specific emotion or some affective state. Distinct emotions are thus defined as distinct patterns of the appraisal variable values (see Figure 6). For example, fear is the felt experience of highly-focused attention, negative valence, and high uncertainty about the outcome and the agent's ability to successfully cope with the situation, along with the physiological components of heightened arousal (K.R Scherer, 2001). Note how changing the value of just one appraisal variable (e.g., coping ability) can change one emotion (fear) into another (anger) (refer to Figure 6).

Given the number of appraisal variables offered by the componential appraisal theories, the resulting space of possible emotions is vast. It therefore enables the representation of a broad range of emotions, including states of varying intensity (e.g., mild annoyance, moderate frustration, anger, rage). It also enables the representation of a wide range of mixed states, and "intermediate or transitional states between the named categories of emotions; with vacillation between emotions that corresponds to uncertain or vacillating appraisals" (Ellsworth & Scherer, 2003).

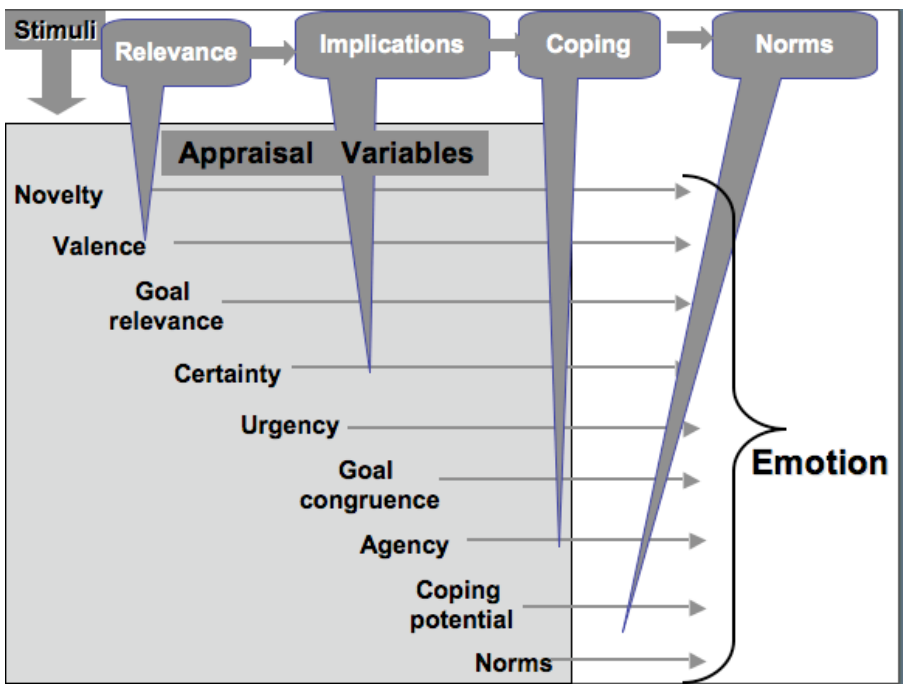
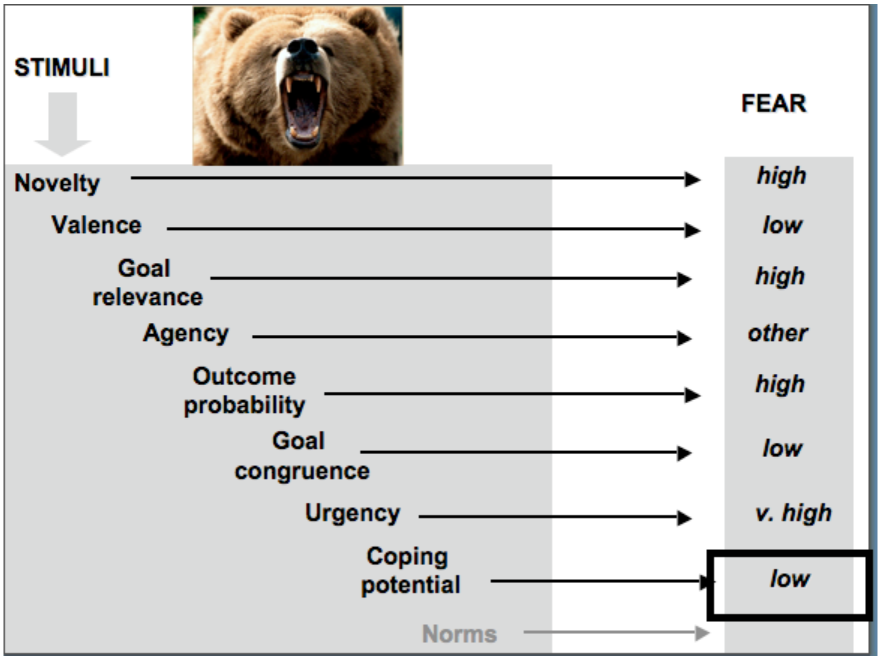


Figure 5: Schematic View of the Componential Theory of Emotion
The figure shows the relationship of the individual appraisal dimensions to the broader categories of evaluations taking place during appraisal (Relevance, Implications, etc.)



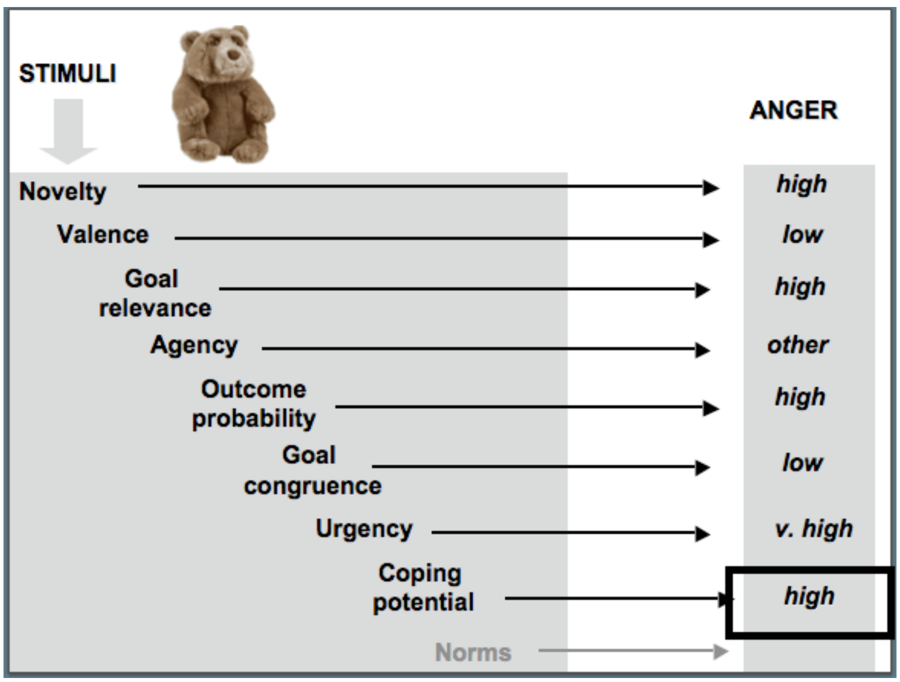


Figure 6: Example of Appraisal Dimension Vectors and Resulting Emotions

The figure illustrates how a change in the value of a single appraisal dimension can make a dramatic difference in the resulting emotion. The upper figure shows the low coping potential, and the resulting fear. The lower figure shows how a change in the coping potential from low to high results in anger, rather than fear.

Table 5: Scherer's Definition of the Mapping of Individual Appraisal Values Onto Specific Emotions

(Based on Table 5-4, pp. 114-115 in (K.R Scherer, 2001)

(nat: natural forces; diss: dissonant; int: intentional; chan: chance; neglig: negligence; obstr: obstruct)

Appraisal Variable	Fear	Anger	Joy	Sadness	Shame	Guilt	Pride
Relevance							
Novelty							
Suddenness	HIGH	HIGH	HIGH/ MED	LOW	LOW	open	open
Familiarity	LOW	LOW	open	LOW	open	open	open
Predictability	LOW	LOW	LOW	open	open	open	open
Valence	LOW	open	open	open	open	open	open
Goal relevance	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH
Implications							
Cause: Agent	OTHER/NAT	OTHER	open	open	self	self	self
Cause: Motive	open	INT	INT/ CHAN	INT/ CHAN	INT/ NEGLIG .	INT	INT
Outcome probability	HIGH	V. HIGH	V. HIGH	V. HIGH	V. HIGH	V. HIGH	V. HIGH
Discrepancy from expectation	DISS	DISS	open	open	open	open	open
Conduciveness to goal	OBSTR	OBSTR	V. HIGH	OBSTR	open	HIGH	HIGH
Urgency	V. HIGH	HIGH	LOW	LOW	HIGH	MED	LOW
Coping Potential							
Control	OPEN	HIGH	open	V. LOW	open	open	open
Power	V. LOW	HIGH	open	V. LOW	open	open	open
Adjustment	LOW	HIGH	MED	MED	MED	MED	HIGH
Normative significance							
Internal standards	open	open	open	open	V,LOW	V.LOW	V.HIGH
External standards	open	LOW	open	open	open	V.LOW	HIGH

3.2.1.3 Comparing OCC with Componential Appraisal Theories

While the two theories of cognitive appraisal described above are distinct, from a computational perspective they share a number of structural similarities. In much of the existing literature, these theories are often discussed separately, which tends to obscure their structural similarities. Below we highlight these similarities, but also point out the differences, and their implications for modeling (refer to

Table 6).

Both theories define emotions in terms of a set of domain-independent features, within an abstract domain (refer to Figure 11 and Table 7). These abstract features represent an intermediate point in the mapping between the stimuli and the emotions proper. For both theories, the specific values of these domain-independent features or variables reflect an evaluation of the emotion eliciting stimuli (e.g., valence, desirability), or the stimuli-agent relationship (e.g., coping potential). OCC refers to these evaluative criteria as the local variables, while the componential models refer to them as the appraisal variables (or appraisal dimensions). The local (in OCC), or appraisal variables (in componential appraisal theories), then represent the agent's *appraisal state,* and correspond to specific emotion.

The set of appraisal variables offered by the componential model is slightly larger, but both sets cover the essential components of appraisal: the valence of the triggers; and their relevance for, and congruence with, the agent's concerns and specific goals. One difference between the two theories is the emphasis on coping in the componential appraisal models, which assign a dedicated appraisal variable to coping.

In the OCC theory, different local variables are associated with different categories of emotions. The componential theories don't make strong assumptions about distinct sets of appraisal variables defining particular emotion classes, although different emotions naturally emphasize different variables; e.g., the social emotions emphasize the appraisal variables related to norms.

The OCC model is more highly structured in terms of the categories of triggers (events, acts by other agents, objects), categories of internal evaluative criteria (goals, behavioral norms, and preferences and tastes), and resulting categories of emotions. This degree of structure is helpful for modeling, since it suggests specific representational structures. Indeed, the detailed specification of emotions in terms of these structures, which map directly onto computational representations, was a key reason for adopting the OCC theory for computational models of emotion generation. On the other hand, neither the degree of structure, nor the particular structure offered by the OCC theory, may always be necessary; e.g., a given model may not need to distinguish among events, acts-of-others or objects to derive an emotion.

Componential theories do not go as far as the OCC theory in suggesting different types of triggers and specific internal structures. Instead, they offer a more uniform structure for emotion calculation, in terms of a single set of appraisal variables, applicable to all categories of emotions, albeit, possibly weighing the distinct variables differently for different emotions (e.g., adherence to social norms is more important for shame and guilt than for sadness and joy). The more uniform treatment of emotions offered by the componential theories has the benefit of providing a simpler computational structure implementing the stimulus-to-emotion mapping.

Componential theories emphasize the role of multiple modalities, although the interactions among these are not quantified to an extent that would readily translate to a multi-modal computational model. Nor is it within the current state-of-the-art to construct such a multi-level model.

Both theories address intensity calculation, primarily in qualitative terms, although the OCC theory offers a more elaborate treatment of the intensity calculation task. Neither theory addresses affective dynamics in detail; that is, the emotion decay rates and approaches for combining multiple emotions, although OCC offer a few precedence rules for conflict resolution.

In contrast to the OCC theory, the componential theory provides some discussion of the emotion effects, suggesting that values of different appraisal variables link directly to specific expressive features or effects on cognition (see discussion below). This feature makes the componential theory attractive, since it enables the use of the same theoretical basis for both emotion generation, and for emotion effects models. The OCC model does not address the consequences of emotions in great depth.

Table 6 provides a comparative summary of the major features of the two models.

Selection of the best appraisal theory for a specific modeling effort depends on a number of factors, including how many and what types of emotions need to be modeled; the extent to which the structures offered by a given theory are useful or helpful for a particular application (e.g., the taxonomy of emotions offered by OCC may or may not be useful for a particular modeling effort); whether both emotion generation and emotion effects need to be modeled, and how important it is to have a single theoretical basis for both models. Which of these theories better reflects the nature of emotion generation via appraisal, remains to be determined.

Table 6: Comparison of OCC with the Componential Appraisal Theories

	OCC	Componential Appraisal Theories
# of emotions possible	22 in original model	Large set possible, within the space defined by the appraisal variables
Emotions accommodated	Basic, some complex & social (see **Table 4**)	Very large set of emotions possible in theory (a subset confirmed via empirical studies_
# of evaluation variables / features	~11	~13 (Scherer's set)
Specific evaluation variables	See **Error! Reference source not found.**	See **Error! Reference source not found.**
Elaboration of the required representational structures	Differentiated taxonomy of emotions and some structures offered (e.g., goals, behavior norms)	Few specific suggestions regarding representational structures, though many implied in the descriptions of the appraisal variables
Elaboration of the mediating mechanisms	Some representational structures offered	High-level descriptions; delineation of distinct processes calculating individual appraisal variable values
Basis for intensity calculations	Qualitative descriptions of components influencing intensity; multiple thresholds for different emotions; distinction between potential and activated emotion	Qualitative descriptions of individual variables affecting intensity; discussion of thresholds and sensitization / habituation
Basis for combining multiple emotions	Some indication of which emotions are likely to be more intense, thus dominant	
Consideration of multiple modalities	Primarily cognitive, but include arousal as a global variable influencing intensity	Yes, but appraisal variables and interpretive processes elaborated primarily within cognitive modality
Consideration of emotion effects on cognition, re-appraisal, expression	Acknowledged but not elaborated	Evidence that some appraisal variables map directly onto cognitive effects or expression components

3.2.2 Emotion Effects

For modeling purposes, it is useful to divide emotion effects into two categories: the visible, often dramatic, behavioral and expressive manifestations, and the less visible, but no less dramatic, effects on the internal perceptual and cognitive processes. (Refer to summaries of emotion effects in Table 1.) Theories of emotion effects are not as elaborated as those for emotion generation. While the dominant approach to emotion generation in affective models is via cognitive appraisal, and two established theories are available (OCC and the cognitive appraisal theories within the componential perspective), no such established theories exist to guide models of emotion effects.

One of the reasons for the lack of an integrated theory is the multi-modal nature of emotion effects. A comprehensive theory of emotion effects would need to account for manifestations of emotions across multiple modalities, including the non-observable effects on the physiological substrate that mediates any observed manifestations, as well as the internal effects on cognition. Such a theory would then in essence explain the phenomenon of emotions. While emotion research has progressed tremendously over the past few decades, it does not yet provide a "unified theory of emotions". In addition, given the diversity of processes mediating emotions, it is unlikely that such a unified theory can ever be established. Given the emphasis on symbolic models and the cognitive modality in existing affective models, we focus below on theories available for modeling emotion effects on cognition.

In most existing affective models, the effects of emotions on cognition are not represented in any depth. Rather, the effects of emotions are manifested via the agent's expressive channels, and behavioral choices. At first glance, modeling the 'invisible' effects of emotions on cognitive processes may seem as an 'overkill' for affective models in models agent architectures. It may seem that a direct mapping of affective states onto their expressive manifestations, and behavioral choices, is adequate. However, much as introducing emotion as the mediating variable between stimuli and responses allows for more flexible mapping between the environment and the agent's behavior, so does the explicit representation of cognitive processes, as the intervening variables between emotions and their expressive and behavioral manifestations. Modeling affective biases on cognition also provides an efficient means of generating appropriate behaviors in more complex environments, and a means of generating personality-dependent affective variability that makes agents more believable.

Several theories have been proposed to explain a particular observed effect of emotions on cognition; e.g., mood congruent recall (Bower, 1981); influence of emotions on judgment and decision-making (e.g. Forgas' AIM (Forgas, 1999); influence of emotions on different cognitive and perceptual processes (Derrybery & Reed, 2003). The theories emphasize different components of information processing (e.g., attention, memory, automatic vs. controlled processes) and researchers often group affective influences into different categories, based on the cognitive structures and processes that are affected. For example, Forgas (Forgas, 2003), focusing on emotion influences on attitudes and social judgments, suggests a distinction between memory-based influences and inference-based influences. An example of the former being network theories of affect, explaining mood congruent recall via spreading activation mechanisms (Bower, 1981). Example of the latter being Schwartz and Clore's theory of affect-as-information (Schwarz & Clore, 1988). Derryberry and Reed (2003), focusing on personality and individual differences research, propose 4 categories of mechanisms whereby emotions influence cognition: automatic activation, response-related interoceptive information, arousal, and attention.

In existing computational models, two theories have been explored: *spreading activation theories* across semantic network representation of memory, and *parameter-based theories,* which suggest that affective factors act as parameters inducing variabilities in cognitive processes (and, subsequently, behavior). Spreading activation theories aim to explain *affective priming* (shorter response times required for identifying targets that are affect-congruent with the priming stimulus vs. those that have a different affective tone), and *mood-congruent recall* (the tendency to preferentially recall schemas from memory whose affective tone matches that of the current mood) (e.g., (Bower, 1992; Derryberry, 1988). Bower's "Network Theory of Affect" assumes a semantic net representation of long-term memory, where nodes representing declarative information co-exist with nodes representing specific emotions. Activation from a triggered emotion spreads to connected nodes, increasing their activation, thereby facilitating the recall of their information. Alternative versions of this theory place the emotion-induced activation external to the semantic net.

A number of researchers have independently proposed a broader theory of mechanisms mediating emotion-cognition interaction, where parameters encoding various affective factors (states and traits), influence a broad range of cognitive processes and structures (e.g., (Hudlicka, 1998; Matthews & Harley, 1993; Ortony et al., 2005; Ritter & Avramides, 2000). The parameters modify characteristics of fundamental cognitive processes (e.g., attention and working memory speed, capacity, and biasing), thereby inducing effects on higher cognition (problem-solving, decision-making, planning, as well as appraisal processes). Several recent models of emotion effects use some variation of this approach (Belavkin & Ritter, 2004; Broekens, Kosters, & Verbeek, 2007; Hudlicka, 2003; Hudlicka, 2007; Ritter, Reifers, Klein, & Schoelles, 2007; Sehaba, Sabouret, & Corruble, 2007). We briefly illustrate the parameter-based approach below, in the context of a cognitive-affective architecture whose focus is the modeling of a broad range of individual differences, including affective states and traits, on cognitive processing: the MAMID architecture (Hudlicka, 1998; Hudlicka, 2002; Hudlicka, 2007; Hudlicka, 2008a). MAMID models emotion effects using a generic methodology for modeling multiple, interacting individual differences, both *stable traits and dynamic states*. Its focus is on modeling the effects of emotions (joy, fear, anger, and sadness) on the cognitive processes mediating decision-making (attention, situation assessment, expectation generation, goal management and action selection). A high-level schematic of the MAMID cognitive-affective architecture is shown in the right part of

Figure 7.

A set of parameters is defined that controls processing within the MAMID architecture modules, for example, the speed and capacities of the different modules, as well as the ranking of the individual constructs processed by these modules (e.g., cues, situations, goals), as they map the inputs (perceptual cues) onto the outputs (selected actions).

Figure 7 illustrates this modeling approach. A particular configuration of emotion intensities and trait values is mapped onto specific values of the architecture parameter values, which then modify the modules' processing (e.g., lower/increase the modules' capacity and speed, introduce a bias for particular types of constructs, such as high-threat or self-related constructs (refer to

Figure 7 and 8). Functions implementing these mappings are constructed on the basis of the available empirical data. For example, the anxiety-linked bias to preferentially attend to threatening cues, and interpret situations as threatening, is modeled in MAMID by ranking high-threat cues and situations more highly, thereby making their processing by the Attention and Situation Assessment modules more likely (refer to

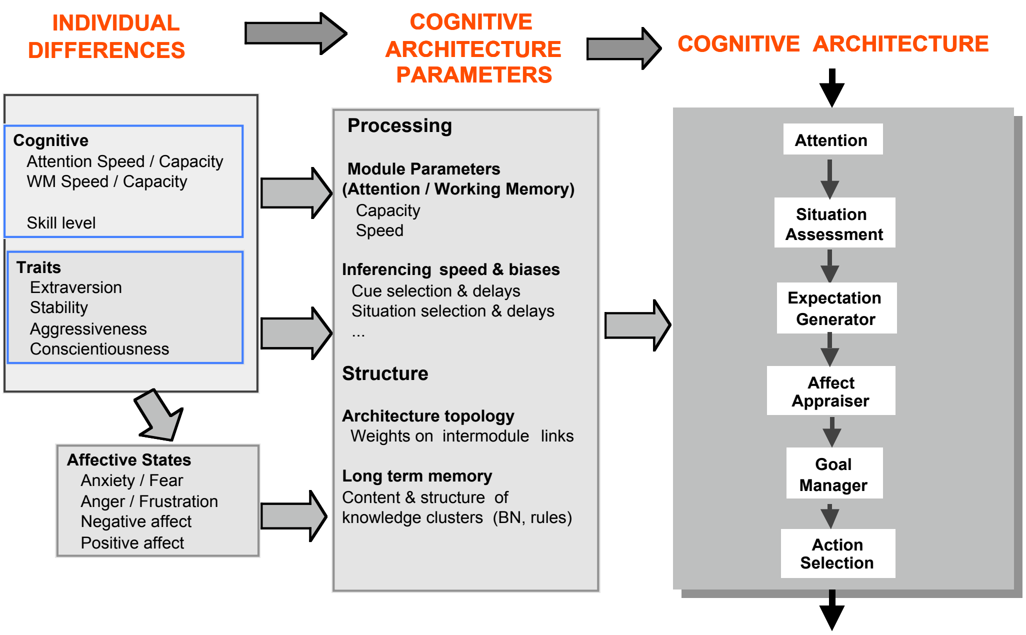
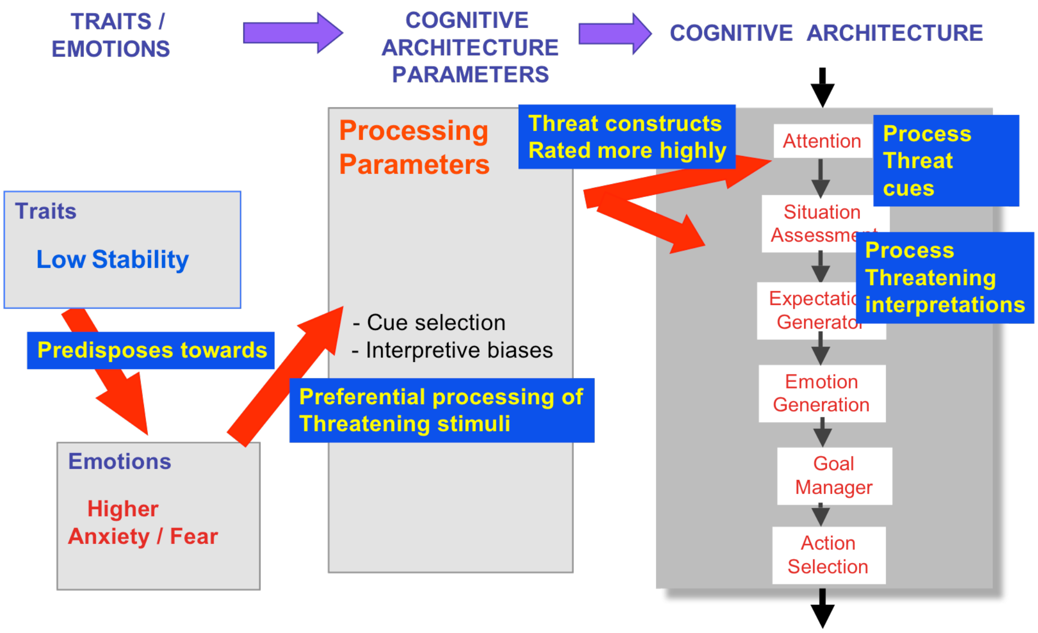


Figure 8). Currently, the parameter-calculating functions consist of weighted linear combinations of the factors that influence each parameter. For example, working memory capacity reflects a normalized weighted sum of emotion intensities, trait values, baseline capacity, and skill level.

The parameter-based models are consistent with recent neuroscience theories, suggesting that emotion effects on cognition are implemented in the brain in terms of



systemic, global effects on multiple brain structures, via distinct patterns of neuromodulation, corresponding to different emotions (Fellous, 2004).

Figure 7: Schematic Illustration of MAMID Methodology Modeling Effects of Individual Differences on Cognitive Processing

Figure 8: Modeling Trait (Low Stability) and State (Anxiety) Threat Bias Within MAMID

3.3 From Theories to Models

Before we discuss the design of affective models in more detail it is helpful to summarize the type of information we would ideally expect existing theories from psychology to provide, to facilitate model development. This pragmatic perspective will then allow us to evaluate candidate theories, to determine which is most appropriate for a given modeling effort.

Ideally, the theories would provide sufficient details to operationalize the computational tasks necessary, and provide specifications of the underlying mechanisms, in terms of representational structures and processes (see also (Broekens, DeGroot, & Kosters, 2008)). Thus, cognitive appraisal theories of emotion generation ought to be able to provide answers to questions such as:

- What is the stimulus-to-emotion mapping for the domain of interest? Should this mapping be implemented *directly* (domain stimuli-to-emotions), or *indirectly*, via some intermediate, domain-independent features (e.g. novelty, valence, desirability, likelihood, responsible agent)?
- How are the influences of external stimuli integrated with internal stimuli (recalled or anticipated events and situations) in triggering emotions?
- What are the distinct stages of the appraisal process and which functions are implemented in each stage? Do these stages vary as a function of the specific emotion, individual, or context?
- What are the dependencies and interactions among the distinct processes implementing appraisal?
- What factors influence emotion intensity, and how?
- What is the nature of the 'affective dynamics', that is, the onset and decay rates of emotions? How do these vary by individuals, emotions and contexts?
- Can multiple emotions be generated by the appraisal process? If not, how should potentially conflicting triggers be integrated into a single emotion? If so, how should these multiple emotions be integrated into a coherent affective state?
- What cognitive structures are necessary to support appraisal (e.g. goals, expectations, plans), and what is the nature and complexity of these structures?
- What levels of complexity are necessary for a particular modeling application (e.g., sensorimotor, schematic, or conceptual)?
- What should be the structure of the emotion construct itself, that is, the object generated by the appraisal process? What information should be represented in this structure (e.g., type, intensity, its triggers, responsible agent, direction, goals involved)?

Theories of emotion effects on cognition ought to be able to provide answers to questions such as:

- Which cognitive processes and structures are influenced by particular emotions, moods, affective states and traits? What is the nature of this influence? What are the effects on dynamic mental constructs such as situations, goals, expectations, and plans?
- How are contents and organization of long-term memory structures affected?
- How is cognitive appraisal affected by emotions?
- What is the relationship between the emotion or mood intensity and the type and magnitude of the influence? Can distinct intensities of emotions or moods have qualitatively different effects on cognitive processes?
- Are distinct emotions the mediating variables of the effects (e.g., fear influences attentional bias towards threat), or are individual appraisal dimensions the mediating variables (e.g., (Lerner & Tiedens, 2006).
- How and when are the influences of multiple emotions, moods and traits combined?
- Are there distinct types of processes that mediate these influences? What are the interactions and dependencies among these processes?
- Last, but not least: Can we obtain sufficient data about these internal processes and structures to enable construction of computational models?

Unfortunately, the existing theories do not provide answers to all of these questions, with the theories attempting to explain the mechanisms of emotion effects on cognition less-well elaborated than those for cognitive appraisal. A given theory may not provide answers at the level of detail necessary for the construction of a computational model. In fact, it is frequently the act of model construction itself that motivates more refinements of the associated psychological theories. While for some of the questions there is significant consensus (e.g., which types of stimuli trigger which emotions), others require considerable 'educated guesswork'; e.g., how opposing emotions should be combined). The least-well defined aspect of emotion generation and emotion effects regards the affective dynamics: the calculation of emotion intensity and emotion effects magnitudes, their changes over time, as well as the integration of multiple emotions and moods, and multiple effects. Typically, only qualitative descriptions of these relationships are available in the psychology literature, although recent research in the expression and recognition of emotions via facial expressions is yielding promising quantitative data regarding facial expression dynamics (e.g., (Cohn, Ambadar, & Ekman, 2007)).

The questions above serve three roles. *First*, they provide a basis for evaluating candidate theories. *Second,* they provide a basis for defining the computational tasks necessary to implement affective models. *Finally*, they help identify specific aspects of affective processing where more concrete theories need to be developed. Below we discuss the computational tasks necessary to implement affective models.

4.0 Computational Perspective on Emotion Modeling

Above, we discussed the theoretical foundations for developing models of emotion generation and emotion effects. We emphasized the theories of cognitive appraisal in emotion generation, as the most useful theoretical basis for symbolic models of emotion. In emotion effects, we discussed several theories of emotion effects on cognition.

Currently, no established, systematic guidelines exist for developing computational models of emotion. We hope to contribute to the development of such guidelines, by delineating the fundamental categories of processes required to model emotions (those modeling emotion generation, and those modeling emotion effects), and by defining the abstract computational tasks necessary to implement these processes.

Below we discuss the core processes and computational tasks in more detail, along with their representational and reasoning requirements.

4.1 Core Processes Required to Model Emotions

In spite of the progress in emotion research over the past 20 years, emotions remain an elusive phenomenon. While some underlying circuitry has been elucidated for some emotions (e.g., amygdala-mediated processing of threatening stimuli, the role of orbitofrontal cortex in emotion regulation), much remains unknown about the mechanisms of emotions. Given the multiple-modalities of emotion, the complexity of the cross-modal interactions, and the fact that affective processes exist at multiple levels of aggregation, it may seem futile, at best, to speak of 'fundamental processes of emotions'.

Nevertheless, for purposes of developing symbolic models of emotions, and modeling emotions in symbolic agent architectures, it is useful to cast the emotion modeling problem in terms of two broad categories of processes. Those responsible for the *generation of emotions*, and those which then mediate the *effects of the activated emotions* on cognition, expressive behavior (e.g., facial expressions, speech) and action selection (refer to Figure 9).

This temporally-based categorization (before and after the 'felt' emotion) provides a useful perspective for computational affective modeling, and helps manage the complexity of the modeling effort, by supporting a systematic deconstruction of these high-level processes into their underlying computational tasks, as discussed below. There are of course many complex interactions among these categories of processes, which must also be represented in order to develop valid models of emotions.

We are not suggesting at this point that these processes correspond to distinct, discrete neural processing mechanisms, but rather that they represent a useful means for managing the complexity associated with symbolic affective modeling. Specifically, this core process view provides a framework that helps organize existing theories and data, as well as a natural hierarchical structure that links emotions roles to the underlying computational tasks necessary to implement them.

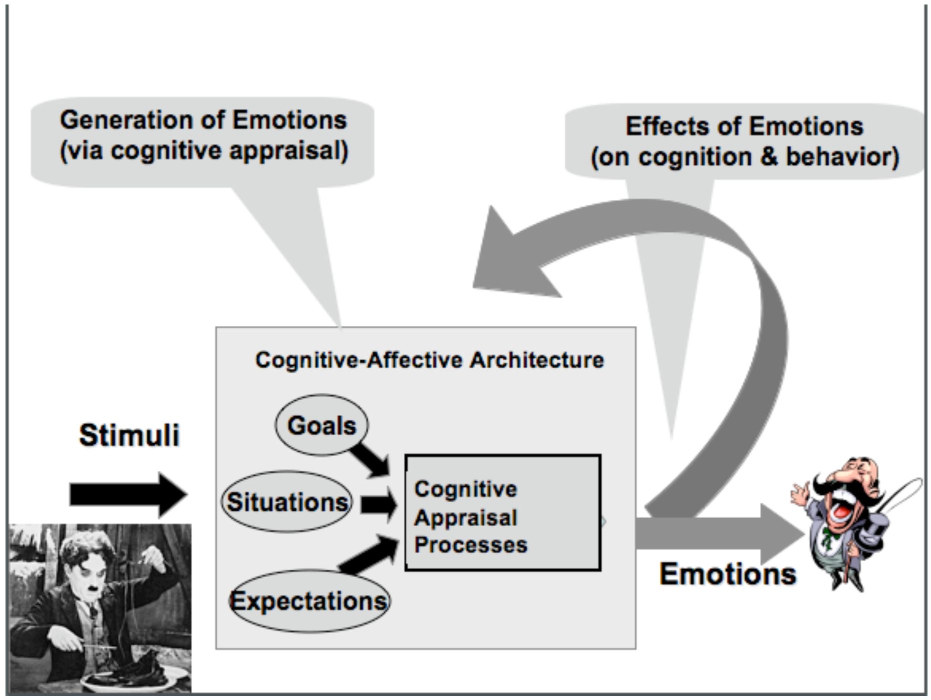


Figure 9: Core Processes of Emotion Necessary for Implementing Symbolic Models of Emotion

4.2 Computational-Task Based Analysis of Emotion Modeling

Identifying the fundamental processes required to model emotions is only the first step. For this perspective to be useful, we must deconstruct these high level processes, and identify the individual computational tasks required to implement emotion generation and emotion effects models.

The objective here is to move beyond the approach where individual models are the organizing dimension, as is typically the case in existing literature on affective modeling, and towards a more general approach, organized in terms of the individual computational tasks, and the associated representational and reasoning requirements. The generic computational tasks required to implement emotion generation and emotion effects discussed below, and further elaborated in section 5, provide a basis for managing the complexity of affective modeling, by providing foundations for more concrete guidelines for model design. They represent generic computational building blocks, from which specific models can be constructed. Figure 10 illustrates the relationship between the computational tasks (bottom part of figure), the fundamental processes (middle part of figure), and the different roles of emotions.

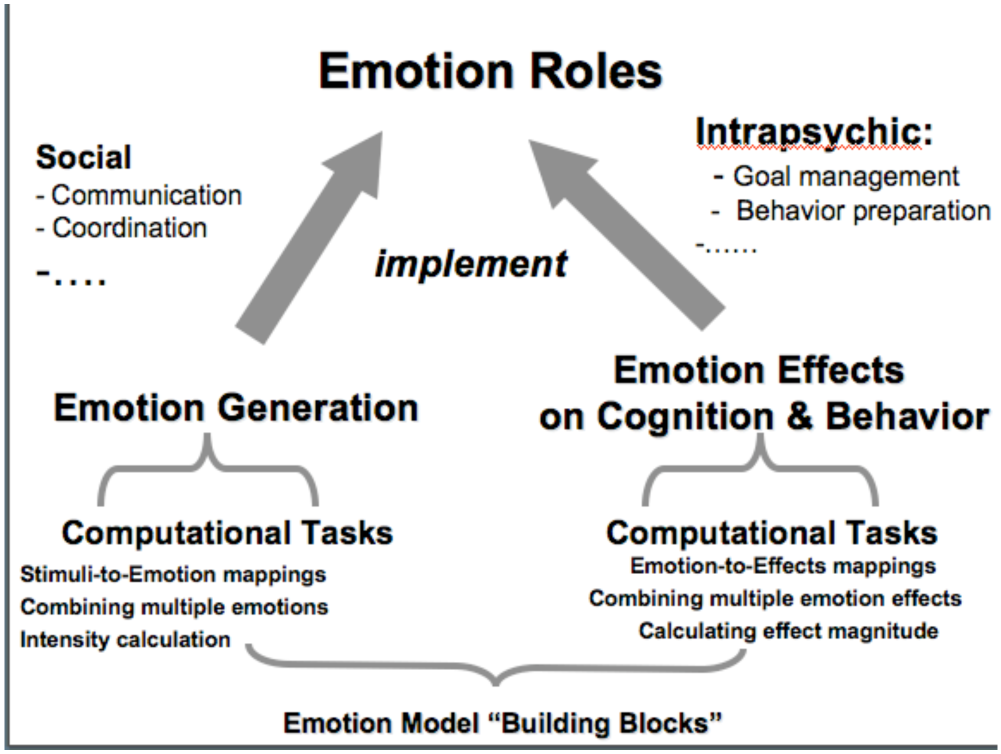


Figure 10: Relationship Between Emotion Roles, the Core Processes of Emotion and the Computational Tasks Necessary to Implement These Processes

4.3 Distinct Domains Required for Modeling Emotions

The high-level computational tasks outlined above provide a basis for defining modeling guidelines, and managing the complexity of affective modeling, but they provide only the first step in the top-down deconstruction of modeling requirements. Another example of a computational and design-oriented perspective on emotion generation is the recent work by Broekens and colleagues, who developed a generic set-theoretic formalism, and an abstract framework, for representing, and comparing, appraisal theories.

Building on the work of Reisenzein (Reisenzein, 2001), Broekens and colleagues (Broekens et al., 2008) offer a high-level, set-theoretic formalism that depicts the abstract structure of the appraisal process, and represents both the processes involved, and the data manipulated. We have augmented their original framework to also represent modeling of emotion effects. The resulting abstract structure is shown in Figure 11.

The framework illustrates the distinct processes involved in emotion generation and emotion effects modeling, and the data manipulated by these processes (e.g., perception (evaluative processes produce a series of mental objects), appraisal (processes that extract the appraisal variable values from the mental objects), and mediation (processes that map the appraisal values onto the resulting emotion(s)). The distinct processes operate on different groups of data, their associated *domains*. Table 7 summarizes these domains and provides examples of their elements: the associated semantic primitives.

This framework thus complements the computational task-based perspective with a set of domains required to implement both emotion generation and emotion effects, and helps define the constituent elements of these domains. These definitions then form a basis for defining the mappings among these domains, necessary to implement emotion generation and emotion effects.

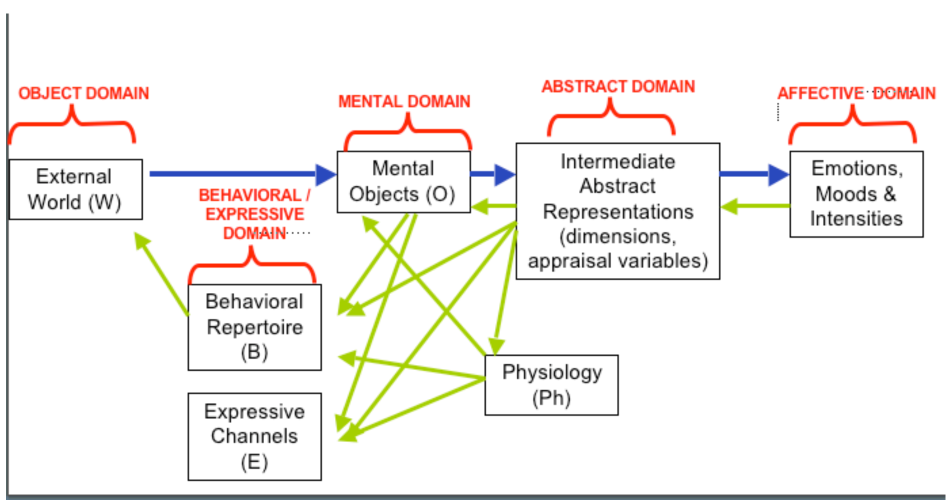


Figure 11: Abstract Framework Representing the Distinct Domains Involved in Emotion Modeling (Augmented Version of an Framework Proposed by Broekens et al., 2008)

Note that this figure assumes the existence of some intermediate, abstract structures reflecting variables that mediate both emotion generation and emotion effects. Not all *affective models necessarily require such an abstract domain. The blue arrows indicate paths mediating emotion generation; the green arrows paths mediating emotion effects.*

Table 7: Domains Required to Implement Affective Models in Agents
The set of domains in the table represents a superset of possible domains required for emotion modeling. For a given model, and given theoretical foundations, only a subset of these may be necessary.

Domain Name	Description	Examples of Domain Elements
Object (W)	Elements of the external world (physical, social), represented by cues (agent's perceptual input)	Other agents, Events, Physical objects
Cognitive (C)	Internal mental constructs necessary to generate emotions, or manifest their influences on cognition	Cues, Situations, Goals, Beliefs, Expectations, Norms, Preferences, Attitudes, Plans
Abstract (Ab)	Theory-dependent; e.g., dimensions, appraisal variables, OCC evaluative criteria	Pleasure, Arousal, Dominance; Certainty, Goal Relevance, Goal Congruence…
Affective (A)	Affective states (emotions, moods) & personality traits	Joy, sadness, fear, anger, pride, envy, jealousy; extraversion
Physiology (Ph)	Simulated physiological characteristics	Level of energy
Expressive Channels (Ex)	Channels within which agent's emotions can be manifested: facial expressions, gestures, posture, gaze & head movement, movement, speech	Facial expressions (smile, frown), speech (sad, excited) gestures (smooth, clumsy), movement (fast, slow) (represented via channel-specific primitives, e.g., FACS)
Behavioral (B)	Agent's behavioral repertoire in its physical & social environment	Walk, run, stand still, pick up object, shake hands w/ another agent

4.4 Generic Computational Tasks

We now discuss the distinct computational tasks necessary to model emotion generation and emotion effects, with an emphasis on the cognitive modality, which is emphasized in symbolic computational models of emotion, and is typically used in agent architectures.

4.4.1 Computational Tasks Required for Modeling Emotion Generation

Figure 12 illustrates a number of distinct computational tasks necessary to implement *emotion generation via cognitive appraisal*, specifically:

- Define and implement the emotion elicitor–to–emotion mapping (depending on the theoretical perspective adopted, this may involve additional subtasks that map the emotion elicitor onto the intermediate representation (PAD dimensions or appraisal variable vectors), and the subsequent mapping of these onto the final emotion(s)).

- Calculate the intensity of the resulting emotion.
- Calculate the decay of the emotion over time.
- Integrate multiple emotions, if multiple emotions were generated.
- Integrate the newly-generated emotion with existing emotion(s) or moods.

Table 8 summarizes the computational tasks required to implement affective dynamics, that is calculation of emotion intensity, its onset and decay, and integration of multiple emotions.

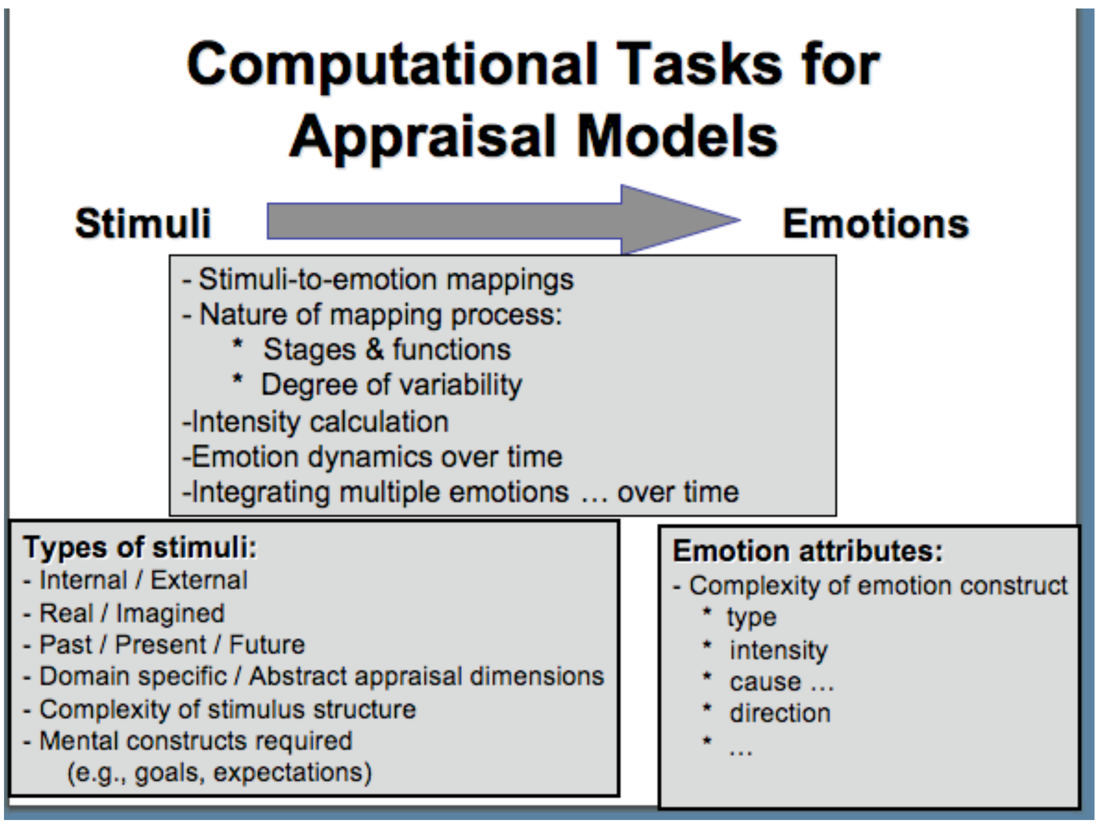


Figure 12: A Computational Perspective on Emotion Generation via Cognitive Appraisal

High-level view of the emotion generation process, showing the inputs (triggering stimuli), output (emotions), and the distinct computational tasks required to implement the stimulus-to-emotion mapping.

Table 8: Summary of Computational Tasks Required to Implement Affective Dynamics in Emotion Effects Modeling

Affective Dynamics Tasks	Examples of Design Alternatives
Intensity Calculation	Step function (0 or 1); Desirability * Likelihood (of situation, event or relevant goal); Modify desirability * likelihood to capture asymmetry in positive vs. negative emotions
Intensity Onset / Decay	Step function, duration for time t; Linear, exponential or logarithmic monotonically increasing / decreasing function
Integrating multiple affective states (similar)	Sum, Maximum, Average, logarithmic, sigmoidal
Integrating multiple affective states (dissimilar)	Max, Precedence rules, Mood congruent emotion selected

4.4.2 Computational Tasks Required for Modeling Emotion Effects

While models of emotion generation typically focus on only one modality (the cognitive modality and cognitive appraisal), models of emotion effects cannot as easily ignore the multi-modal nature of emotion. This is particularly the case in models implemented in the context of embodied agents that need to manifest emotions not only via behavioral choices, but also via expressive manifestations within the channels available in their particular embodiment (e.g., facial expressions, gestures, posture etc.)

The multi-modal nature of emotion effects increases both the number and the type of computational tasks necessary to model emotion effects. Figure 13 illustrates the abstract computational tasks required to model the *effects of emotions* across multiple modalities, specifically:

- Define and implement the emotion/mood–to–effects mappings, for the modalities included in the model (e.g., cognitive, expressive, behavioral, neurophysiological). Depending on the theoretical perspective adopted, this may involve additional subtasks that implement any intermediate steps, and are defined in terms of more abstract semantic primitives provided by the theory (e.g., dimensions, appraisal variables).
- Determine the magnitude of the resulting effect(s) as a function of the emotion or mood intensities.
- Determine the changes in these effects as the emotion or mood intensity decays over time.

- Integrate effects of multiple emotions, moods, or some emotion and mood combinations, if multiple emotions and moods were generated, at the appropriate stage of processing.
- Integrate the effects of the newly-generated emotion with any residual, on-going effects, to ensure believable transitions among states over time.
- Account for variability in the above by both the intensity of the affective state, and by the specific personality of the modeled agent.
- Coordinate the visible manifestations of emotion effects across multiple channels and modalities within a single time frame, to ensure believable manifestations.

Table 9 summarizes the computational tasks required to implement affective dynamics, that is calculation of emotion intensity, its onset and decay, and integration of multiple emotions.

The specific tasks necessary for a particular model depend on the selected theoretical perspective; e.g., discrete/categorical models don't require a two-stage mapping sequence, and the specific theory of emotion generation or emotion effects that is implemented in the model. Note also that not all models necessarily require all tasks; e.g., in simpler models, where only one emotion is generated, there is no need to integrating multiple emotions. These tasks, and associated design choices, are discussed in more detail in section 5.

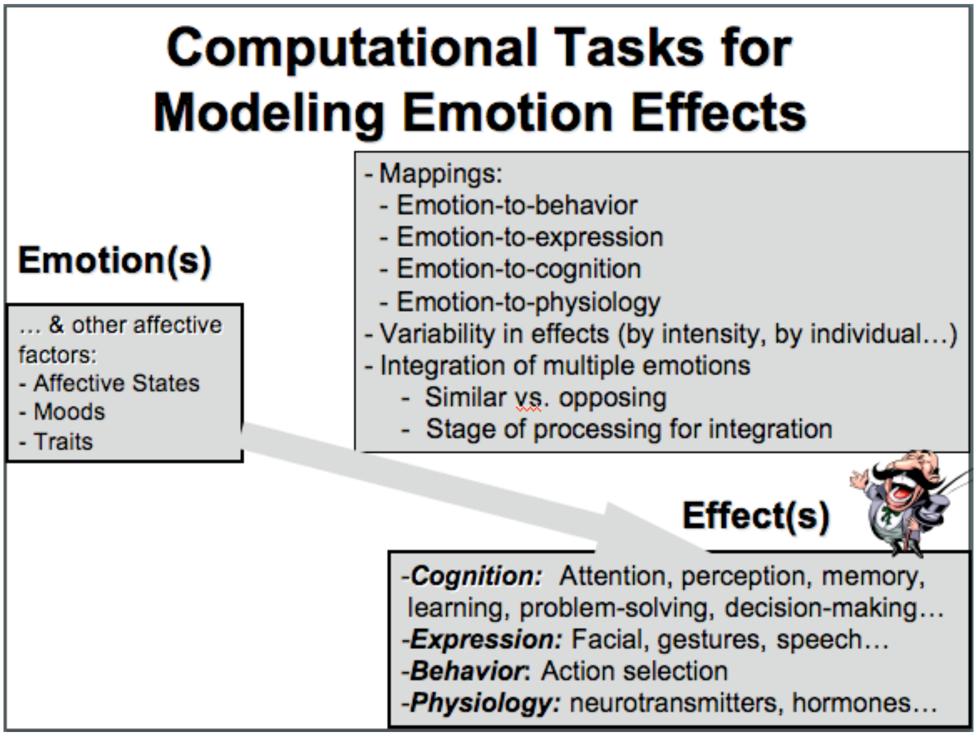


Figure 13: Computational Tasks Necessary to Model Emotion and Mood Effects

Table 9: Summary of Computational Tasks Required to Implement Affective Dynamics in Emotion Effects Modeling

Affective Dynamics Tasks	Examples of Design Alternatives
Effect Magnitude Calculation	In expressive domain: larger effect in individual expressive elements (size of smile); more expressive elements used; duration In behavioral domain: distinct actions associated with different intensities of emotion
Effect Magnitude Onset / Decay	In expressive domain: Empirically based onset, peak, offset values for individual elements (e.g., facial muscles)
Integration of multiple effects	Max (select effects associated with emotion with highest intensity); combine intensities of individual elements within an abstract domain, map results onto effects (PAD dimensions, appraisal variables)

4.5 Representational and Reasoning Requirements

The diagrams in Figure 11, Figure 12 and Figure 13 provide a temporal perspective on emotion modeling. When we begin to discuss the representational requirements, we also need to consider an abstraction-hierarchy perspective, specified at three levels of abstraction: abstract entities, their attributes, and the representational formalisms used to implement them.

At the highest level, are the *abstract entities* that need be represented, to enable the interpretations necessary to implement cognitive appraisal, and to mediate the multi-modal emotion effects. Different types of abstract entities exist in the different domains involved in emotion modeling. For example, in the *object domain,* examples of entities are events, situations, other agents, and physical objects, as well as more complex causal structures. In the *cognitive domain*, examples of these entities are goals, expectations, standards and norms, frequently organized into hierarchies. Specific entities required depend on the types of affective states in the model. For example, complex emotions require reasoning about the self (e.g., comparing own attributes or behavior with others, and with established norms), thus requiring an explicit representation of the self. In the *abstract domain*, the entities are theory-specific abstractions, such as the OCC evaluation criteria and the componential theories' appraisal variables. In the *affective domain,* the entities are the structures that represent the affective states derived by the emotion generation process: emotions, moods or undifferentiated positive and negative affect. In the domains within which emotion effects are manifested, the entities are different for the distinct modalities and within the different expressive channels. Table 10 and

Table 11 provide examples of different entity types required for emotion generation and emotion effects modeling, respectively.

These entities are associated with different attributes that contain the information necessary to implement emotion generation and effects modeling. For example, a goal entity may need attributes to represent the following information: type (e.g., active pursuit vs. interest vs. replenishment (in the OCC goal taxonomy), importance, difficulty, likelihood_of_success, objects/agents required to achieve the goal, subgoals/parent_goals, status (e.g., in_progress, completed, failed), etc. A more in-depth discussion of these attributes is beyond the scope of this paper but can be found in (Hudlicka, forthcoming).

Once the high-level specifications outlined above are defined, the actual representational formalism must be selected. We focus here on symbolic formalisms, since the high-level models of emotion discussed here typically use symbolic representations (vs. connectionist or analytical models). A number of established representational formalisms are available, including: Bayesian belief nets, production rules, frames, predicate calculus, fuzzy logic formalisms, as well as procedural knowledge, encoded in terms of demons, knowledge-sources, or arbitrary functions.

The choice of a specific formalism is guided by several factors, including the need for an explicit representation of uncertainty, the ability to represent temporal information, compatibility with the domain and the mental constructs required, ability to effectively represent the semantic primitives offered by the selected theory (e.g., appraisal variables or OCC evaluative criteria), ability to implement learning and generalization. There are also pragmatic considerations, which include compatibility with an existing system, availability, and familiarity personal preferences.

The complexity of reasoning required to implement the affective processes depends on the complexity of the domain, the types of affective states represented in the model, and the modalities and channels within which emotions can be expressed. A significant domain-specific inferencing may be required to interpret the events and situations in the object domain, in terms of the evaluation criteria provided by the selected theory (e.g., OCC evaluative criteria, appraisal dimensions). For any but the most simple domains, this will require representation and reasoning under uncertainty. Different emotions then require different degrees of reasoning complexity. For example, the OCC prospect-based emotions, that is, emotions involving expectations of future events or situations, require complex causal representations of the object domain, and temporal reasoning, such as what-if and abductive reasoning. (Note, however, that for the simple forms of the basic emotions the interpretive processes are innate and 'hardwired' (e.g., fear of snakes; large, looming objects, etc.).

Table 10: Examples of Abstract Entity Types Required for Emotion Generation

Object Domain	Mental Domain	Abstract Domain	Affective Domain
Agents, Events, Physical objects, Causal diagrams	Cues, Situations, Goals, Expectations, Norms & Standards,Preferences, Plans	Interpretive features (OCC evaluation criteria, appraisal variables); Mood dimensions (pleasure, arousal, dominance)	Emotions, Moods, Affect

Table 11: Examples of Abstract Entity Types Required for Modeling Emotion Effects

Behavioral Domain	Physiological Domain
Different action types	Arousal, Hormone
Expressive Domain	**Cognitive/Mental Domain**
Facial expression (eye brows, lip corners,…); Speech (pitch, pitch range, volume, rate); Gesture (limb, joints)…; posture (lean forward/back, arms open/closed)	Cues, Goals, Plans, Beliefs, Situations, Expectations…

5.0 Guidelines for Emotion Model Design

We now describe an approach to emotion model design that is based on the computational tasks outlined above, which implement the two core affective processes: emotion generation and emotion effects. We present the design process in the context of designing an agent architecture, thereby focusing on the development of more applied affective models, aimed a enhancing the agent's autonomy and affective realism. (As opposed to models designed to elucidate the underlying mechanisms of affective processes.)

Given the limited space available, we describe in detail only a subset of the design steps. A full description of the design process can be found in a forthcoming book on affective computing (Hudlicka, forthcoming).

To identify the degree to which emotions need to be modeled in a given agent, and to select the most appropriate approach to affective modeling, the model designer first needs to conduct a requirements analysis: s/he needs to have a clear understanding of how explicit modeling of emotion in the agent architecture will enhance its functionality or effectiveness. Since this requirements analysis focuses on the emotions and their roles, we refer to it as *affective requirements analysis* and *affective design process*. To maintain, and reinforce, the computational perspective emphasized in this paper, we emphasize the use of the distinct domains and abstract computational tasks mediating emotion generation and emotion effects, which were introduced earlier. Table 12 provides a summary of the steps required for the affective requirements analysis and model design process.

Table 12: Summary of the Steps Comprising Affective Requirements Analysis

STEP	Description	Example of Results
#1: Define functions of emotions in game & within each agent	Define specific interpersonal & intrapsychic roles of emotions.	Enhance autonomy via emotion-based behavioral routines (avoid danger by enabling fear); Support cooperative behavior via attachment; Enhance realism.
#2: Define specific emotions & moods required	Define elements of the affective domain (A) needed to implement the roles identified in step 1	Emotions: joy, sadness, fear, anger, boredom, surprise, disgust Moods: happy, sad, fearful, jealous
#3: Define emotion triggers (in physical & social environments; self). Define agent's behavioral repertoire.	Define elements of the object domain (W) & Behavior domain (B),	W: approaching user, smiling agent, physical object in the environment B: smile at user, pick up object.
#4: Define agent's embodiment, including any simulated physiology	Define expressive channels (face, gestures, speech)-Expressive Domain (Ex). Define 'physiological' variables (arousal, energy)-Physiological domain (Ph)	Face: express joy, sadness, fear, anger, boredom, surprise Speech: content & prosody for emotions above. Arousal: define arousal levels corresponding to different emotions.
#5: Define level of model abstraction.	Define degree to which internal mental & cognitive states will be modeled – Cognitive domain (C)	Interpretations of current situation, Expectations, Goals, Plans
#6: Select theoretical perspective & specific theory for emotion generation & emotion effects modeling.	Theory determines nature of Abstract domain (Ab), if any.	If componential: Ab consists of appraisal variables; If OCC: Ab consists of the OCC evaluation criteria (local variables); If dimensional: Ab consists PAD
#7: Define emotion trigger & emotion effects associations	Define mappings among domains defined in Stp. 3-6 (emotion generation & emotion effects). If an abstract domain is used, define mappings via abstract primitives.	Cues-to-Emotion: Friendly user→Joy; Evil monster (in game agents) → Fear; Irritating user → Anger. Emotion-to-Cognition: Fear → reduced attentional focus, bias toward threatening cues and interpretations Emotion-to-Behavior: Fear--> Run (in NPC agents); Anger → Use rude speech to user;
#8: Define affective dynamics	Define intensity functions for emotion/mood generation; Emotion integration; Emotion effects magnitude & integration of multiple effects.	Intensity = desirability * likelihood; Intensity decays exponentially Winner-take-all conflict resolution for conflicting emotions

STEP	Description	Example of Results
#9: Specify Abstract Computational Tasks Required	Define abstract computational tasks necessary to implement generation of emotions and their effects, including affective dynamics, if any.	Triggers-to-Emotion Mapping; Trigger to Appraisal Variables Mapping & App. Variable to Emotion Mapping; Emotion to Cognitive Effects Mappings; Intensity calculating functions
#10: Select representational & reasoning formalisms	Select most appropriate representational & inferencing formalisms	Rules, Bayesian belief nets, Fuzzy logic

5.1 Define Objectives and Roles of Emotions: Why Should the Agent Architecture Include Emotions?

Emotions may be needed to enhance the *agent's affective and social realism*, thereby its believability, and ultimately its effectiveness when interacting with a human user. Emotions may also be included to enhance the *agent's autonomy* and *effectiveness*, particularly in more complex and uncertain environments, by biasing the agent's behavior towards a selection of particular goals and actions. Such biasing helps prune the large problem search spaces that can be generated in complex, uncertain environments. Emotions can make the agents more autonomous, and produce more adaptive behavior; e.g., an agent with appropriate levels of fear will successfully evade dangers and survive longer in hostile environments. Note, however, that just because these capabilities can be mediated by emotions does not mean that affective mechanisms represent the only way to implement such capabilities.

Defining the objectives allows a definition of the specific roles that emotions should perform in the agent. Table 13 provides a summary of possible roles of emotions in agents, for the distinct objectives outlined above.

Table 13: Affective Requirements Analysis Step #1: Possible Roles of Emotions in Agents

Increased Character Autonomy	
Intrapsychic Emotion Roles	
Homeostatic Functions	Ensures that agent maintains internal resources to survive & thrive in its environment (virtual or physical)
Motivational Functions	Ensures agent avoids specific dangers, seeks out specific rewards, exhibits appropriate assertiveness/aggressiveness to accomplish goals & obtain resources, enables learning and seeking of information via boredom & curiosity
Control of behavior	(Re) prioritize goals as necessary in response to events, implement processing biases to help coordinate cognition & behavior, rapid (re)allocation of resources to address a new situation / event, rapid detection of salient stimuli
Interpersonal Emotion Roles	
Coordination and communication of behavioral intent	Facilitate more realistic behavior toward other agents and the users, by communicating internal state & intentions, communicate social status information
Attachment behavior	Allows agents to form relationships with one another, creating more effective collaborations, and more interesting interaction with the user, manage & repair relationships by communicating social norms
Increased Character Affective Realism (Believability)	
Intrapsychic Emotion Roles	
Homeostatic Functions	Agent's resource level enables them to exhibit more realistic behavior (e.g., grumpy, irritable when tired; happy when rested).
Motivational Functions	Different reactions to the same situation/event, as a function of changing moods & goals
Interpersonal Emotion Roles	
Coordination and communication of behavioral intent	Communicate territorial rights, social goals (form relationships, social status)
Attachment / friendships	Communicate different types of emotions with different agents and users, depending on the nature of the relationship

5.2 Define Specific Emotions Needed to Implement Selected Roles

There are many possible emotions, varying in complexity of triggers and manifestations. (Clore and colleagues have identified over 200 emotions terms by analyzing a set of 600 English terms with an affective component (G.L. Clore & Ortony, 1988). In addition, there are moods, mixed states with both cognitive and affective components (e.g., attitudes), and affective personality traits. A given application cannot represent all of these states. The designer must therefore choose which emotions will be represented. The majority of existing models represent a subset of the basic emotions, that is: joy, sadness, anger, fear, disgust and often surprise. Researchers are increasingly attempting to represent more complex emotions. For example, Becker-Asano has used the dimensional theoretical perspective to represent the emotions of hope, relief and fears-confirmed in terms of the PAD dimensions (Becker-Asano, 2008). The number and types of emotions required has implications for the selection of a theoretical perspective, as we shall see below.

5.3 Define Emotion-Eliciting Triggers and Agent's Behavioral Repertoire

This step involves an analysis of the agent's environment, either virtual or physical, and identification of its perceptual inputs that can trigger emotions. This represents the agent's input cues, and the elements of the object domain (W). The agent's possible behaviors are also defined at this stage, which represent its behavioral repertoire, and define the elements of the behavioral domain (B). The elements comprising the W and B domains are necessarily domain-specific, that is, specific for the agent's particular physical and social environment.

5.4 Define Agent's Embodiment

The nature of the agent's embodiment depends on a number of factors, including the its overall objectives (e.g., a game character, a virtual coach, a robot used for rehabilitation, a robotic companion for children, etc.), on the function of the emotions in the agent (e.g., autonomy vs. affective realism), the overall visual realism of the context (e.g., an application for a mobile device with a small screen vs. a life-size agent image in a virtual reality setting vs. a robot); and the resources available for developing the agent's embodiments (e.g., detailed facial expressions, speech); and the emotions which the agent needs to express.

The nature of the agent's embodiment then determines which expressive channels are available for manifesting its emotions. Recall that the expressive channels within which emotions can be manifested are: facial expressions, gaze, head movement, gestures, body posture, nature of character's movement in general, and speech (both verbal content and non-verbal sounds, as well as prosodic speech qualities such as pitch, rate, and volume). The designer must determine which of these expressive channels are necessary for the agent's objectives, and which are feasible to implement, given any software and hardware constraints. The designer also needs to consider which expressive channels are best for manifesting the emotions that are to be included in the model, given any implementation constraints, such as screen size. Typically, facial expressions are best for expressing emotions. However, in situations where the screen size is small, and its resolution low, other channels may be more appropriate (e.g., speech, gestures).

The designer must also take into consideration the state-of-the-art in generating affective expressions in a given channel; e.g., established representational and animation techniques are available for display affective expressions on the face, most using the elements of the Facial Action Coding System (P. Ekman & Friesen, 1978) as the semantic primitives, whereas affective synthetic speech is not yet as advanced.

Defining the agent's embodiment in effect defines the individual channels within the expressive domain (Ex), and established their semantic primitives.

Another component of the agent's embodiment is its simulated physiology. Not all agents require an explicit model of the physiological modality. However, given the embodied nature of most agents, some representation of some simulated physiological characteristics will make the characters appear more realistic, by enabling the simulated physiological states to influence agent behavior. This is the case both for modeling the contribution of these characteristics to emotion generation (e.g., a fatigued or hungry agent may be more likely to become irritated or angry), and for modeling the effects of emotions (e.g., a game NPC experiencing prolonged fear may become more rapidly fatigued, and increased fatigue may make the NPC more irritable). To the extent that physiological characteristics are modeled in affective agents, they are highly abstracted and simple, typically consisting of some measure of energy (e.g., energy, arousal), and available physical resources (e.g., degree of hunger, fatigue and 'health'). For example, representing arousal in agent Max (Becker, Nakasone, Prendinger, Ishizuka, & Wachsmuth, 2005), allows him to display varying rates of breathing and blinking, thereby enhancing its believability. These features then represent the elements of the *physiological domain* (Ph) of the affective model.

The end result of this step in the design process is therefore a definition of the expressive (Ex) and physiological (Ph) domains.

5.5 Define Model's Level of Abstraction

The model designer must next decide on the degree to which internal mental states and constructs should be explicitly represented in the model. This in effect defines the model's level of abstraction, which ranges between two extremes: black-box models, which do not require detailed representations of internal mental states, process-level models, which aim to simulate or emulate the underlying mechanisms, and represent constructs in the *cognitive domain* (C), such as beliefs, goals, expectations, plans etc. These constructs often suggest which modules need to be represented in the agent architecture (e.g.. Goal Manager, Planner), and defining the cognitive domain therefore begins to define the structure of the agent architecture.

Black-box models implement direct mappings between emotion triggers and the resulting emotions, and the emotions and the resulting behavior. (In many models, there are in fact intermediate representations, in terms of abstract features such as the OCC evaluative criteria or the componential theory appraisal variables, but these models do not necessarily explicitly represent cognitive processes.)

In contrast, process-level models aim to explicitly represent the cognitive processing involved in generating emotions, and the affective biases on cognition that influence decision-making, planning and action selection. This indirect mapping, enabled by the explicit representation of the cognitive domain, then enables a greater degree of differentiation and variability in both processes (emotion generation and emotion effects), which is not possible in black-box models.

Figure 14 illustrates the differences between a black-box model (top part) and a process-level model (bottom part), in the context of emotion effects modeling. The blue notes indicate reasons why one or the other alternative might be used. The figure also includes a third alternative, a deeper model that does not aim to emulate the actual biological mechanisms, and is used to facilitate coordination among effects across multiple modalities.

For any but the most simple models, the agent will need an explicit representation of goals, to enable dynamic and realistic generation emotions. More complex agents may require an explicit representation of additional mental constructs, such as perceptions, beliefs, expectations, and plans. These constructs then form the elements of the *cognitive domain (C)*. No fixed set of established semantic primitives exists in the cognitive domain, beyond the high-level constructs listed above and in Table 7, but literature in cognitive, social and emotion psychology is a good source of potential additional high-level constructs (e.g., attitudes).

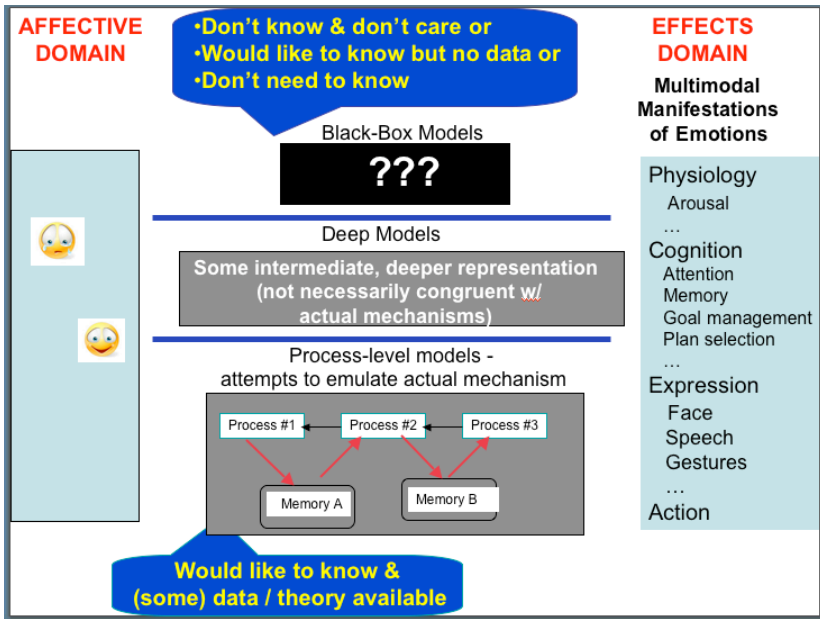


Figure 14: Alternatives in the Levels of Model Abstraction

5.6 Select Theoretical Foundations for the Affective Model

We have already emphasized the importance of modeling emotion generation and emotion effects in a theoretically and empirically grounded manner. By this we mean that the affective models should use an established *theoretical perspective* on emotions (e.g., discrete, dimensional, componential), and, where available, use an *established theory* to support the modeling of a particular process (e.g., OCC cognitive appraisal theories for emotion generation). Both theoretical perspectives, and the theories supporting emotion generation and emotion effects modeling, were discuss in section 3.

The theoretical perspectives differ in the semantic primitives they offer for representing emotions (basic emotions in the discrete / categorical perspective, pleasure-arousal or pleasure-arousal-dominance in the dimensional perspective, and the appraisal variables in the componential perspective). They also differ in the degree to which they define, and require, an abstract domain to mediate the trigger-to-emotion mappings and emotion-to-effects mappings. In the discrete / categorical perspective, no intermediate, abstract features are necessary. In the others, the dimensions or the appraisal variables provide these domain-independent features, which represent the elements of the abstract domain. For some of these abstract features, data exist that support a direct mapping from specific feature values to effects in the behavioral or expressive domain; e.g., novelty induces raising of brows and eyelids, interruption or inhalation in voice, straightening of posture (see Scherer (Klaus R. Scherer, 1992), table 5.3 for a list of appraisal variable-effects associations.)

For many others, these data are not available. Thus in some cases the same set of abstract features can serve both to mediate emotion generation and emotions effects. The perspectives also differ in the number of emotions they can accommodate; i.e., the space defined by the PA or PAD dimensions is smaller than that defined by the appraisal variables. Finally, the perspectives differ in the degree to which its semantic primitives support manifestations of emotions in the expressive channels, and combination of multiple emotions. For example, the dimensions in the dimensional perspective are well-suited to support continuous and gradual modification of facial expressions, which is necessary for affectively-realistic facial expressions. In contrast, use of the discrete / categorical perspective to support facial expression modeling requires additional computation to implement affectively-realistic affective dynamics.

The differences outlined above help determine which perspective is best suited for a particular agent. For example, if the agent has only a few of the basic emotions, and does not need to represent both fear and anger (or any other pairs of emotions characterized by the same values of arousal and valence), then the dimensional perspective, with just two dimensions (arousal and valence (aka pleasure)), is appropriate. If both fear and anger are required, then a dimensional perspective may still be appropriate, but a third dimension needs to be added (dominance), to distinguish between fear and anger, both characterized by negative valence and high arousal. If a large number of emotions is required, including complex and social emotions, the componential perspective is more appropriate, since it can accommodate a large number of affective states. The discrete / categorical perspective could be used for either of the above cases, but would require an explicit addition of emotion-specific processes, or modules for each emotion represented in the model.

Figure 15 illustrates the different theoretical perspectives in modeling emotion generation, and

Figure 16 illustrates these perspectives in the context of emotion effects modeling.

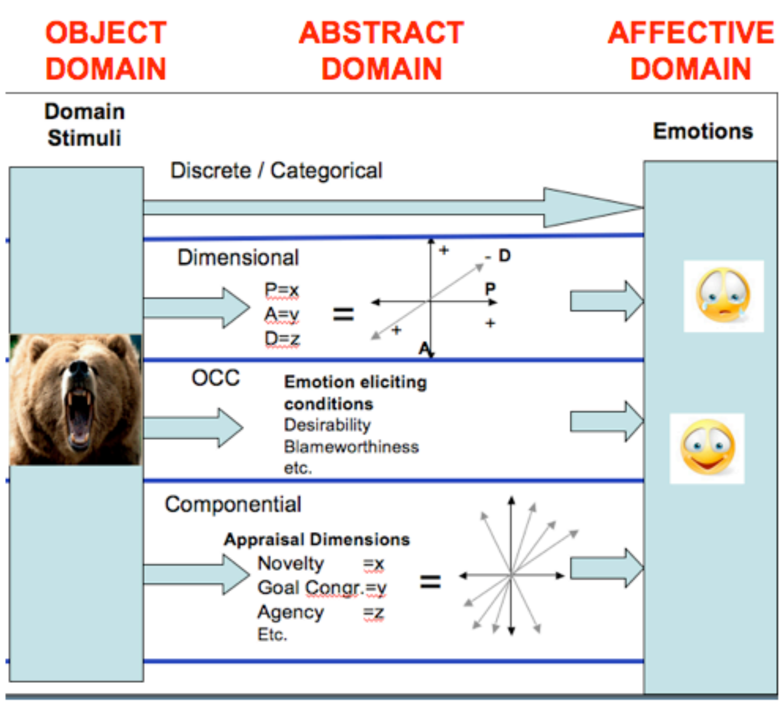


Figure 15: Alternative Theoretical Perspectives on Emotion Generation

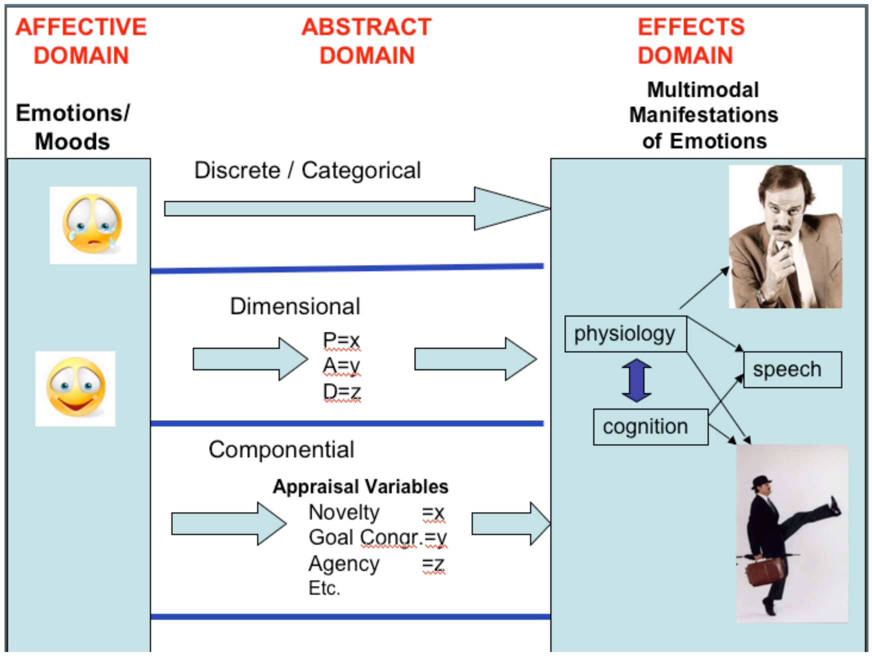


Figure 16: Alternative Theoretical Perspectives on Emotion Effects Modeling

We now turn to the specific theories available to support *emotion generation* and *emotion effects modeling*. The most elaborated theory for *emotion generation* is cognitive appraisal, and this is also the most frequent basis for modeling emotion generation in affective agent models. The dominant theory in cognitive appraisal is the OCC theory, which provides an elaborate taxonomy of emotion triggers, as well as domain-independent evaluative criteria used to generate different emotions. More recently, theories of emotion generation within the componential perspective have begun to be used as basis of affective modeling, which also emphasize the use of domain-independent evaluative criteria, termed the *appraisal variables.* Section 3 describes these theories and the associated domain-independent evaluative criteria in more detail.

Table 6 above compares the OCC and the componential appraisal theories.

The OCC theory has the advantage of having been implemented numerous times, and the existing models of emotion generation can therefore serve as concrete examples for developing analogous models in agents. The disadvantage of the OCC theory is that the OCC abstract evaluative criteria (e.g., pleased about an event undesirable for another, disapproving of own behavior) don't readily map onto emotion effects. If the OCC theory is selected as the basis of an agent emotion generation, some other theory (or approach) must be selected as a basis of modeling emotion effects.

The componential theory's appraisal variables have the advantage that they can serve both to support emotion generation, and emotion effects modeling, since many of the appraisal variables values map onto elements within the expressive domain. Recently, efforts are being made to also identify mappings between appraisal variables and the cognitive domain (e.g., high certainty appraisal appears correlated with the use of heuristic processing, vs. analytical processing, in problem solving).

Regardless of the theory selected, the abstract elements provided by the theory then serve to define the *abstract domain* (Ab). In the case of OCC, the domain-independent evaluative criteria represent the individual elements of the abstract domain (refer to Table 4 for a complete list of these evaluative criteria). In the case of the componential theory's appraisal variables, the individual appraisal variables are the elements of the abstract domain (refer to Table 5 for a list of appraisal variables).

Neither theory elaborates the role of non-cognitive triggers in emotion generation in great detail, beyond acknowledging their existence. In agents that also represent a simulated physiological modality, the emotion generation should take the values of particular physiological characteristics into account, to ensure the agent's affective realism and believability.

Due to the multimodal nature of emotion effects, and in contrast to emotion generation, no dominant theory exists supporting *emotion effects modeling*, which would be analogous to the OCC theory in emotion generation. The consequence of this lack of an all-encompassing theory to support modeling of the cross-modal effects of emotions is that different theories must be used for the different modalities. We differentiate here between the internal effects of emotions on cognition, and the external manifestations of emotions across the expressive channels, and in behavior.

Regarding the effects of emotions on cognition (assuming cognitive processing is represented in the agent): the parameter-based theories lend themselves well to support modeling of the diverse effects of emotions on cognition. A range of parameters is defined to control distinct aspects of processing within the affective model and the agent architecture. For example, parameters may control the speed and capacity of attention, situation assessment, goal management, planning, and even emotion generation itself. When an emotion is generated in the agent, it is translated into a set of specific parameter values, which then induce changes in the architecture modules. The resulting modifications of the cognitive processes are eventually reflected in decisions made by the agent, and the specific actions selected. Refer to Figure 7 and 8 for an illustration of this approach.

Regarding the effects of emotions on the expressive channels, the diversity of processes mediating the different effects across the distinct channels (e.g., facial expression, posture, gestures, speech), and the lack of a 'unified theory' of effects, would suggest that the effects of emotions may best be represented by sets of direct mappings, across the expressive channels available in the agent's embodiment.

However, in spite of the lack of a 'unified theory' of emotion effects, it is often beneficial to have an underlying deeper representation of information processing, which is not necessarily a process-level model, but which facilitates coordination among the multiple modalities within which emotions are manifested, and provides a more robust and flexible basis for modeling than a collection of direct emotion-to-effect mappings (refer to Figure 14). This point has been made by a number emotion modelers in the context of architectures for believable embodied agents (Aylett, 2004; de Rosis et al., 2003; Prendinger, Descamps, & Ishizuka, 2002).

5.7 Define Specific Triggers and Effects of Emotions Represented in the Model

In this step the mappings mediating emotion generation and emotion effects are defined. The former involves defining the mappings that map the emotion eliciting triggers to particular emotions. The latter involves defining the mappings that map the different emotions and moods onto their manifestations, in the modalities available in the agent. Depending on the theoretical foundations selected for the model, and the nature of the associated abstract domain, if any, some of the mappings may involve the abstract domain. For example, if the dimensional theoretical perspective is used, the triggers may be mapped onto the dimensions, which may then be mapped onto particular emotions. Alternatively, the dimensions may be mapped directly onto particular effects (e.g., high arousal will map onto faster movements, and simulated physiological characteristics such as faster breathing rate).

Defining these mappings consists of two steps. *First,* the domains involved in the mapping are specified (refer to Table 7 for a summary of the domains); e.g., in a model that uses the *discrete / categorical perspective*, where the distinct emotions are the semantic primitives, the mapping mediating emotion generation will map elements from the object domain (W) directly onto the affective domain (A). A model using the *componential perspective* would first involve mapping the object domain elements onto the appraisal variables within the abstract domain (Ab), which would then be mapped onto the affective domain. *Second,* the specific contents of the mappings must be specified, which will implement the generation of specific emotions or moods, and their effects. Minimally, this means effects on behavior, but typically also expressive manifestations and, less often, effects on cognitive processing. Since some agents may also have simulated 'physiological' characteristics, such as energy level, fatigue etc., both the effects of emotions on these characteristics, and the effects of these characteristics on emotion generation, may also need to be represented, to ensure affective realism.

Minimally, the following domains, and their associated elements, are required for an agent affective model: Object domain (W), Affective domain (A), and Behavioral domain (B). Typically, the agents will express the generated emotions, requiring the definition of the Expressive domain (Ex) and its elements. In more complex agents, the Cognitive (C) domain may also need to be defined, to enable the agent to produce more interesting and affectively-realistic behavior, as a result of the generated emotions. In many agents the physiological domain (Ph) is also defined, in terms of characteristics reflecting the agent's 'physical' state (e.g., fatigue, energy). In addition, abstract domains may be defined, which reflect particular theoretical perspectives and/or specific theories, and whose elements provide additional variables that mediate emotion effects across multiple modalities; e.g., parameters used to represent effects of emotions on different cognitive processes.

The specific contents of these mappings should be based on empirical data regarding the typical triggers of particular emotions, and the typical effects of emotions on the modalities and channels represented in the model, including any mediating abstract domains defined, instantiated within the specific context of the agent's world and its functionality. For example, the fact that a physical threat to a robot's safety triggers fear would be instantiated in the robot's context by identifying the specific threats to its safety (e.g., fall off a platform, fall into a ditch, get crushed by a rock). The empirical evidence that fear induces a focus on threatening stimuli, and an interpretive bias to assess neutral situations as dangerous, would then be instantiated in the robot's context by distinguishing among high- and low-threat stimuli (e.g., quickly-approaching large object is a high-threat stimulus), and defining alternative interpretations of situations, along the danger level spectrum, to enable the fearful agent to select the more dangerous ones, and react accordingly.

The result of this step is a set of mappings among the domains represented in the agent, which then serve as a basis for defining the computational tasks necessary to implement the emotion generation and emotion effects in the agent's affective model. The exact set of the mappings necessary depends on the expressive modalities represented in the agent, which are determined by its particular embodiment, and on the abstract domains represented in the model, if any, which are determined by the model's theoretical basis. Table 14 and

Table 15 summarize the results of this phase of the affective requirements analysis and design process.

Note that we assume here that the agents don't learn, and therefore the specific contents of the mappings from the affective to the behavioral domains must be pre-defined by the designer. This constraint is likely to change in the future, as machine learning becomes increasingly incorporated into gaming.

Table 14: Affective Requirements Analysis Step #7: Defining the Mappings Mediating Emotion Generation

Theoretical Perspective	Structure of Required Mappings	Examples of Contents
Discrete/ Categorical	W → A	***Dangerous event*** → fear ***Another agent obstructs goal*** → anger ***Friendly user*** → joy
Dimensional	1. W –> Ab, where Ab is either {pleasure,arousal} or {pleasure, arousal, dominance} 2. Ab → A	***Dangerous event*** → V=low, A=high, D=low → fear ***Agent obstructs goal*** → V=low, A=high, D=high → anger ***Friendly user*** → V=high, A=medium → joy
Componential	W –> Ab, where Ab consists of n-tuples of appraisal variable values 2. Ab → A	***Dangerous event*** → novelty=high, valence=low, goal relevance=high, goal congruence=low, coping=low → fear ***Agent obstructs goal*** → novelty=high, valence=low, goal relevance=high, goal congruence=low, coping=high → anger ***Friendly user*** → novelty=high, valence=high, goal relevance=high, goal congruence=high → joy
OCC	1. W → Ab, where Ab consists of the OCC evaluative criteria 2. Ab → A	***Dangerous event*** → displeased, neg. fortunes-of-self, prospect-based → fear ***Agent obstructs goal*** → displeased, disapproving, neg. fortunes-of-self, negative well being, expectation deviation → anger ***Friendly user*** → pleased, pos. fortunes-of-self, positive well being → joy

Table 15: Affective Requirements Analysis Step #7: Defining the Mappings Mediating Emotion Effects

Theoretical Perspective	Structure of Required Mappings	Examples of Contents
Discrete/ Categorical	A → Effects domains (subset of domains used in a particular agent)	fear → facial expr. of fear, fearful speech, run or hide anger → facial expr. of anger, angry speech, aggressive behavior toward other agent joy → facial expr. of joy, happy speech, approach friendly user
Dimensional	Ab → Effects domains, where Ab is either {valence,arousal} or {valence, arousal, dominance} and / or A → Effects domains	V=high -> approach. V=low+D=low→ avoid, escape. V=low+D=high→ avoid/approach; A=high→ high energy, physio. variables. A=medium -> med. energy in behavior, physio. variables. A=high→ high energy, physio. variables. (Note that multiple dimensions may be required to determine an effect, and that often a single effect cannot be uniquely determined via dimension values alone.)
Componential	Ab → Effects domains, where A consists of n-tuples of appraisal variable values and / or A → Effects	Novelty = high→ orient attention; raise eyebrows/lids; interrupt speech; straighten posture; Valence=high→ raise lip corners; approach Valence=low→ lower lip corners; avoid Coping=high→ approach, full voice Coping=low→ avoid or freeze, (Based on Scherer, 1992)

5.8 Define Affective Dynamics

Affective dynamics in emotion modeling refer to calculating emotion intensities, during emotion generation, and determining the magnitudes of emotion effects, during emotion effects modeling, as well as determining the onset and decay of both the intensities and the magnitude of the emotions' effects. Defining the affective dynamics of a particular model also involves deciding how to address the integration of multiple emotions (during emotion generation), and multiple effects (during emotion effects modeling), either when multiple affective states are generated, or when attempting to integrate newly generated state(s) with existing ones.

In section 4 we discussed the alternative approaches to modeling affective dynamics. As already pointed out, realistic representation of affective dynamics represents a major challenge in affective modeling, with few guidelines and fewer underlying theories available, as well as a paucity of empirical data, beyond statements of qualitative relationships (e.g., anxiety biases attention towards detection of threatening stimuli; arousal induces higher pitch and rate of speech). In many existing affective models quantification of the available qualitative data often ad hoc, with specific values determined empirically, based on model performance.

Calculations of emotion intensity typically involves simple formulas (e.g., emotion intensity = desirability of some event x its likelihood), or linear combinations of differentially-weighted contributing factors. These formulas may require significant fine-tuning to adjust model performance. Specifying the emotion intensity calculation function therefore involves selection of the function itself, as well as the variables used, and their relative weights. In other words, selecting the features of the elements of the object domain, or the abstract domain, and their weights.

The simples design choice for emotion intensity calculation would be to represent intensity as a binary value (0 or 1), with the intensity being 1, if the triggers of the affective state were present, 0 otherwise. However, for agents requiring higher levels of affective realism, varying degrees of intensity must be represented. Typically, intensity values range between 0 and 1 (or some other fixed interval), with different combinations of triggering stimuli, and their characteristics, resulting in different intensity values. Table 16 summarizes the most commonly used functions for emotion intensity calculation, and lists their benefits and drawbacks.

Table 16: Examples of Intensity Calculating Formulas

Intensity function	Pros / Cons	Model / Person Using the Function
Importance * belief (that event is true or that goal will be affected)	+ Simple + Explicit representation of agent's belief + Works for many simple cases - Ignores asymmetry in success/failure of goal - Ignores expectation of event - Ignores possible differences between actual likelihood and agent's beliefs	De Rosis et al., 2003 (Greta)
(Desirability x (change in) (Likelihood of success) (for positive emotions) |Undesirability| x (change in) (Likelihood of failure) (for negative emotions)	+ Relatively simple + Accounts for change in perceived likelihood of success/failure + Works for many simple cases Captures asymmetry in success or failure of affected goal - Requires distinct variables for importance of success (goal desirability)(joy & hope) vs. Importance of avoiding failure (goal undesirability) (distress & fear)	Reilly, 1996 (Em)
|desirability| * likelihood	+ Simple + Works for many simple cases - Ignores asymmetry in success/failure of goal - Ignores expectation of event - Ignores changes in likelihood	Gratch & Marsella, 2004 (EMA)
(1.7 * desirability * expectation**.5) + (-.7 * desirability) (for positive emotions) (2 * desirability * expectation**2) – desirability (for negative emotions)	+ Explicitly represents asymmetry in importance of success vs. avoiding failure - Constants are empirically derived and likely to be context specific	El Nasr et al., 2000 (FLAME)

The next step in affective dynamics model design, is to decide how the emotion intensity should change over time, that is, the rates of the emotion onset and decay. The simplest choice here is to eliminate this component altogether, and use a simple step function, with the emotion simply appearing in its full intensity, lasting for some time interval, and then returning to zero or its baseline value for that character. Different temporal intervals may be defined for different emotions, and for different agent personalities (e.g., a generally happy agent (high extraversion, high agreeableness, and low neuroticism in terms of the Five Factor model traits) will maintain positive emotions for longer time periods; an irascible agent will become angry quickly and will remain angry for longer time periods).

Alternatively, the onset and decay rates may follow some monotonically increasing or decreasing (respectively) function, to model the affective dynamics in a more realistic manner. A variety of functions have been used in emotion generation models, including linear, exponential, sigmoid and logarithmic (Reilly 2006). Table 17 summarizes the most frequently used functions for modeling emotion intensity decay.

Table 17: Alternatives for Modeling Emotion Intensity Decay

Function Type	Descriptions	Pros	Cons
Linear	Decrement intensity at t-1 by a decay constant	Simple to compute	Not realistic
Exponential	Decrement at each t is proportional to intensity at t-1; slope determined by decay constant; faster than logarithmic	More realistic than linear	
Logarithmic	Decrement at each t is proportional to intensity at t-1; slope determined by decay constant; slower than exponential	More realistic than linear	
Mass spring	Decrement at each t is proportional to intensity at t-1; slope determined by decay constants; sinusoid behavior	More realistic than linear & exponential for modeling arousal & valence decay (Reisenzein, 1994)	More complex computational required

Finally, the designer must decide how multiple emotions should be integrated. Multiple emotions may be generated by the appraisal process, and existing emotion(s) must be combined with newly-generated emotion(s). This aspect of affective dynamics is least well developed, both in existing psychological theories and conceptual models, and in computational models. Typically, relatively simple approaches are used to address this complex problem, which limits the realism of the resulting models in any but the most simple situations..

Reilly has analyzed several existing approaches to combining similar emotions (positive with positive, negative with negative), and highlights their drawbacks and benefits, as follows. Simple addition of intensities can lead to too much intensity (e.g., few 'low intensity' emotions lead to a 'high intensity' reaction). Averaging the intensities may result in a final intensity that is lower than one of the constituent intensities: an unlikely situation in biological agents. Max (or winner-take-all) approach ignores the cumulative effects of multiple emotions. The limitations of these simple functions have motivated other approaches, including logarithmic and sigmoid functions (Reilly, 2006; Picard, 1997). In many cases, customized, domain-dependent weightings are used, so that a particular emotion is preferentially generated, as a function of the character's personality. For example, high-extraversion characters may be more likely to feel positive emotions, whereas high-neuroticism characters may be more likely to feel negative emotions (Hudlicka, 2007). Table 18 summarizes existing approaches, and lists their benefits and drawbacks.

Table 18: Alternative Formulas for Combining the Intensities of Similar Emotions

Model	Benefits / Drawbacks
Simple addition	- Does not reflect human affective dynamics - Produces too much intensity too fast - A few low-intensity emotions result in a high-intensity emotion
Average	- Does not reflect human affective dynamics - Final result can be less than the most intense component
Max (winner-take-all)	+ Responsive to high-intensity emotions - Ignores cumulative effects of multiple emotions - Does not use all emotions
Sigmoidal function	+ Uses all emotions + Models saturation effect + Closer to human affective dynamics by addressing problems above + Linear at mid-range
Logarithmic function	+ Uses all emotions + Closer to human affective dynamics by addressing problems above + Linear at low-intensity ranges

A more problematic situation occurs when opposing or distinctly different emotions are derived (e.g., a particular situation brings both joy and sadness). Neither the available theories, nor the existing empirical data, currently provide a basis for a principled approach to this problem. Should opposing emotions cancel each other out? (Are we likely to feel calm and neutral if our house burns down but we have just won the lottery?) Is it even appropriate to think of emotions in pairs of opposites? Can we assume that the strongest emotion is the appropriate one, as some models do (e.g., Hudlicka's

MAMID (Hudlicka, 2004; Hudlicka, 2007)? At what stage of processing are emotions combined and any contradictions resolved? Should conflicting emotions be resolved at the appraisal stage, to avoid the problem entirely? At the cognitive effects stage, e.g., during goal selection? Or at the behavior selection stage? (Refer to Figure 17 for an illustration of these alternatives in emotion effects modeling.) The latter being potentially the most problematic; and yet it is apparent that this phenomenon occurs in biological agents. One only needs to witness the scrambling of a frightened squirrel as a car approaches to see a dramatic impact of the failure to resolve contradictory behavioral tendencies.

The computational solutions for combining non-congruent emotions are generally task- or domain- specific, and often ad hoc. Existing approaches fall into one of four categories: *intensity based* (choose emotion or mood with highest intensity); *precedence-rule based* (define precedence relationships for the emotions represented in the agent, and resolve any conflict among multiple states via these rules; precedence relationships are often defined based on the agent's personality); *mood-congruence based approaches* (select an emotion that is congruent with the current mood); and *abstract representation based* new state definition (combine the semantic primitives within which an emotion or mood is represented (e.g., PAD dimensions, appraisal variables), and generate a new affective state that integrates the multiple states). (See Hudlicka (forthcoming) for a more in-depth discussion of this problem.)

Once the emotions are generated, the dynamics of their effects need to be modeled. This process is made more challenging by the multi-modal nature of emotion effects, discussed above, since different approaches are appropriate for the different modalities, or channels.

Modeling different magnitudes of effects in the behavioral domain often involves the use of multiple intensity thresholds for manifestations of different emotions across distinct channels (e.g., different intensity thresholds may be required to trigger emotion effects on cognition vs. facial expression vs. behavior) (Sonnemans & Frijda, 1994). Multiple thresholds are also often used to map different intensities onto distinct actions. For example, different intensities of joy can be reflected in laughing, clapping hands, or jumping up and down. Different intensities of fear can manifest in different speeds of escape, or, in extreme cases, in complete paralysis.

In several of the channels within the *expressive domain,* the emotion intensity is directly proportional to the magnitude of the associated effects, as expressed by the semantic primitives associated with the particular expressive channel. For example, in facial expression, the higher the intensity of happiness, the greater the movement of the zygomatic muscles will be, causing a larger smile (up to a maximum determined by the associated effectors). Similarly, the effects of increasing intensity of anger on the prosodic qualities of speech, will be reflected in higher values of some of the prosodic features, such as volume and rate. Note, however, that there are significant individual differences in the expressive manifestations of emotions, as well as cultural differences. It is important that the model designer take these into account to create believable non-playing characters. For example, in an interactive drama application, agent A might express its anger by shouting very loudly, whereas agent B might suddenly become very calm and controlled, and speak in very low tones.

Intensity can also be expressed by involving an increasing number of the available semantic elements in a given channel, as well as by involving a larger number of expressive channels. For example, within the facial expression channel, increasingly intense smile involves not only larger displacement of the lip corners, but also the involvement of the eyebrows, as well as a longer duration of the displacement. When the intensity increases further, a smile may progress to laughter, which also involves the speech channel and different vocalizations, as well as changes in the speech features that reflect changes in prosody (e.g., pitch, rate, volume).

Regarding the onset and decay of the emotion effects, the choices here are analogous to those for the onset and decay of emotion intensity during emotion generation. However, implementing the onset and decay of emotion and mood effects is more complex, due to the multiple channels across different modalities that are involved in emotion manifestations. The dynamics of multiple processes must be considered here. In many cases, these processes involve physiological systems that must run their course once triggered, and that have associated refractory periods. For example, increased arousal associated with heightened intensity of many emotions (joy, fear, anger) is mediated by changes in concentrations of specific hormones and neurotransmitters, which must be dissipated or neutralized before the system can return to its baseline state. Thus different effects are manifested along different time scales.

Many of these subtleties can be ignored in existing game characters. As was the case with emotion intensity, the simplest choice is to ignore the onset and decay rates altogether, and simply manifest a particular set of effects at a given magnitude over some temporal interval. The next simplest alternative is to translate the changing intensity value directly into the corresponding magnitudes within the effects domains included in a particular agent. In facial expressions, this would mean translating different intensities onto the degree of displacement of specific facial muscles from their neutral position. In speech, this would mean translating different intensities into magnitudes of different speech signal features, such as volume and rate. These translations would need to take into account individual differences in manifesting different emotions by different agents, as a function of their personality.

However, as demands for agents' affective complexity and realism increase, these simple solutions will no longer be adequate and the issues of multiple, independently-evolving processes mediating emotion effects across multiple channels will increasingly need to be addressed by affective model designers. In addition, some affective states can only be distinguished via the distinct patterns of their affective dynamics (e.g., embarrassment). Believable manifestations of these types of complex emotions in embodied agents is an active area of research.

As was the case with intensity calculation, distinct patterns of emotion and mood effects can be defined for distinct agent personalities, to further increase their affective realism and enhance the player's engagement. This can be done by specifying different thresholds for the distinct manifestations of different emotions, thereby emphasizing the expression of a particular emotion, and / or the expression of a particular emotion within particular channels. Personality differences can also be reflected in the onset and decay rates of different effects. For example, agent with personality A may have rapidly-increasing and intense positive reactions, with the positive emotions decaying very slowly, whereas character with personality B might have generate negative emotions with high intensity, and rapid decay rates.

For a more in-depth discussion of this issue, and a discussion of modeling varying magnitudes of emotion effects on cognition, see Hudlicka (forthcoming).

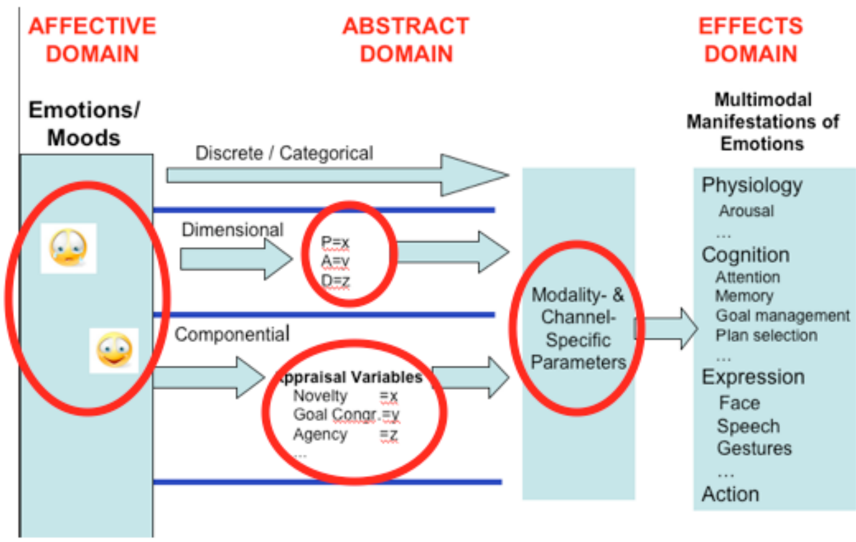


Figure 17: Multiple Points (Marked by Red Circles) Within the Processing Sequence where Multiple Emotions and Moods Can be Integrated

6.0 Summary and Conclusions

Recognizing the lack of consistent terminology and design guidelines in emotion modeling, this paper proposes an computational analytical framework to address this problem. The basic thesis is that emotion phenomena can usefully be understood (and modeled) in terms of two fundamental processes: *emotion generation* and *emotion effects,* and the associated computational tasks. These tasks serve as the computational buildings blocks of affective models, and involve, for both processes the following: defining a set of *domains;* defining a set of *mappings among* these domains (from triggers to emotions in emotion generation, and from emotions to their effects in the case of emotion effects*)*, defining *intensity and magnitude calculation* functions to compute the emotion intensities during generation, and the magnitude of the effects, and functions that *combine and* *integrate* multiple emotions, both similar and dissimilar emotions, and both during emotion generation and emotion effects modeling.

This analysis represents a step toward formalizing emotion modeling, and providing foundations for the development of more systematic design guidelines, and alternatives available for model development.

Identifying the specific computational tasks necessary to implement emotions also helps address critical questions regarding the nature of emotions, and the specific benefits that emotions may provide in synthetic agents and robots.

The analysis presented here has several limitations, partly due to lack of space, but more importantly, due to the fact that the necessary validation and analysis of existing models do not yet exist. *First,* only the cognitive modality of emotion was discussed; both emotion generation via cognitive appraisal, and emotion effects on cognition. This was due both to lack of space and to the predominance of cognitively-based models of emotion, and in no way suggests that the other modalities of emotion are not as critical for understanding these complex phenomena. *Second,* the treatment of the various alternatives for computing the three fundamental computational tasks (mappings, intensity and magnitude, and integration) was necessarily superficial. In part due to lack of space, but primarily because systematic evaluation and validation of existing (or possible) alternatives have not yet been established. *Third*, lack of space did not allow for an exhaustive discussion of existing models, and only brief references were made to existing models. All three of these limitations will be partially addressed in a forthcoming textbook (Hudlicka, forthcoming).

It is hoped that the analysis presented here will stimulate a more focused dialogue about the design of computational affective models, a refinement of the proposed computational analytical framework, and contribute to a more systematic approach to the design of affective models, by specifying both the constituent computational building blocks, and the alternatives available for their implementation.

The focus in this paper has been on designing affective models in agent architectures – thus, a more application-oriented, pragmatic focus. This is in contrast to affective models whose objective is to elucidate the mechanisms underlying affective processes. Such models could not as easily ignore the multi-modal nature of emotion generation, as many existing agent models do, and would need to explicitly represent the complex interactions among the multiple modalities, in both generation and emotion effects modeling.

Development of emotion models in agent architectures is certainly an area where affective modeling can contribute to enhanced performance and effectiveness. However, a no less important function of affective modeling is the ability of these models to enhance our understanding of affective phenomena. Computational modeling necessitates detailed operationalizations of existing high-level constructs, including the very term 'emotion', and thereby has the potential to contribute to the elucidation of the underlying mechanisms of affective processes. This is perhaps the greatest challenge, and promise, of computational affective modeling.

7.0 References

Andre, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (2000). *Exploiting Models of Personality and Emotions to Control the Behavior of Animated Interactive Agents* Paper presented at the Proceedings of IWAI, Siena, Italy.

Arnold, M. B. (1960). *Emotion and personality.* . New York: Columbia University Press.

Averill, J. R. (1994). I Feel, Therefore I Am - I Think. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions*. Oxford: Oxford University Press.

Aylett, R. S. (2004). *Agents and affect: why embodied agents need affective systems 3rd Hellenic Conference on AI, Samos, May 2004 Springer Verlag LNAI 3025 pp.* Paper presented at the 3rd Hellenic Conference on AI, Samos, Greece.

Bates, J., Loyall, A. B., & Reilly, W. S. (1992). Integrating Reactivity, Goals, and Emotion in a Broad Agent, *Proceedings of the 14th Meeting of the Cognitive Science Society*.

Becker, C., Nakasone, A., Prendinger, H., Ishizuka, M., & Wachsmuth, I. (2005). *Physiologically interactive gaming with the 3D agent Max.* Paper presented at the International Workshop on Conversational Informatics at JSAI-05, Kitakyushu, Japan.

Becker-Asano, C. (2008). *WASABI: Affect Simulation for Agents with Believable Interactivity*: IOS Press. .

Belavkin, R. V., & Ritter, F. E. (2004). *OPTIMIST: A new conflict resolution algorithm for ACT-R.* Paper presented at the Proceedings of the Sixth International Conference on Cognitive Modeling, Pittsburgh, PA.

Bower, G. H. (1981). Mood and Memory. *American Psychologist,, 36*, 129-148.

Bower, G. H. (1992). How Might Emotions Affect Memory? In S. A. Christianson (Ed.), *Handbook of Emotion and Memory*. Hillsdale, NJ: Lawrence Erlbaum.

Breazeal, C., & Brooks, R. (2005). Robot Emotion: A Functional Perspective. In J.-M. Fellous & M. A. Arbib (Eds.), *Who Needs Emotions?* NY: Oxford University Press.

Broekens, J., DeGroot, D., & Kosters, W. A. (2008). Formal Models of Appraisal: Theory, Specification, and Computational Model. *Cognitive Systems Research, 9*(3), 173-197.

Broekens, J., Kosters, W., & Verbeek, F. (2007). On Affect and Self-adaptation: Potential Benefits of Valence-Controlled Action-Selection. *Bio-inspired Modeling of Cognitive Tasks*, 357-366.

Cañamero, L. (1997). A hormonal model of emotions for behavior control, *4th European Conference on Artificial Life (ECAL '97)*. Brighton, UK.

Cañamero, L., & Avila-Gracia, O. (2007). *A Bottom-Up Investigation of Emotional Modulation in Competitive Scienarios.* Paper presented at the Affective Computing and Intelligent Interaction.

Cañamero, L. D. (2001). Building Emotional Artifacts in Social Worlds: Challenges and Perspectives, *AAAI Fall Symposium "Emotional and Intelligent II: The Tangled Knot of Social Cognition"* Cape Cod, MA: Menlo Park, CA: AAAI Press.

Clore, G. L. (1994). Why Emotions Are Felt? In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions*. Oxford: Oxford University Press.

Clore, G. L., & Ortony, A. (1988). The Semantics of the Affective Lexicon. In V. Hamilton, G. Bower & N. Frijda (Eds.), *Cognitive science perspectives on emotion and motivation* (pp. 367-397). Amsterdam: Martinus Nijhoff.

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-Based Measurement of Facial Expression with the Facial Action Coding System. In J. A. Coan & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment*. NY: Oxford.

de Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & De Carolis, B. (2003). From Greta's mind to her face: Modelling the dynamics of affective states in a conversational embodied agent. *International journal of human-computer studies, 59*(1-2), 81-118.

Derryberry, D. (1988). Emotional influences on evaluative judgments: Roles of arousal, attention, and spreading activation. *Motivation and Emotion, 12*(1), 23-55.

Derrybery, D., & Reed, M. A. (2003). Information Processing Approaches to Individual Differences in Emotional Reactivity. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. NY: Oxford.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3-4), 169-200.

Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. NY: Oxford.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system* Palo Alto, CA: Consulting Psychologists Press.

El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). FLAME - Fuzzy Logic Adaptive Model of Emotions. *Autonomous Agents and Multi-Agent Systems, 3*(3), 219–257.

Elliot, C. (1992). *The Affective Reasoner: A Process Model of Emotions in a Multiagent System.* Northwestern University, Evanston.

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal Processes in Emotion. In R. J. Davidson, K. R. Scherer & H.H.Goldsmith (Eds.), *Handbook of Affective Sciences*. NY: Oxford University Press.

Fellous, J. M. (2004). *From Human Emotions to Robot Emotions.* Paper presented at the AAAI Spring Symposium: Architectures for Modeling Emotion, Stanford University, CA.

Forgas, J. (1999). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117*(1), 39-66.

Forgas, J. (2003). Affective Influencees on Attitudes and Judgments. In K. R. S. R.J. Davidson, H.H.Goldsmith (Ed.), *Handbook of Affective Sciences*. NY, NY: Oxford University Press.

Frijda, N. (2008). The Psychologists' Point of View. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed.). NY: The Guilford Press.

Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Frijda, N. H. (1994). Emotions Are Functional, Most of the Time. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions*. New York: Oxford University Press.

Frijda, N. H., & Swagerman, J. (1987). Can computers feel? Theory and design of an emotional system. *Cognition and Emotion, 1*(3), 235-257.

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

Hudlicka, E. (1998). *Modeling Emotion in Symbolic Cognitive Architectures.* Paper presented at the AAAI Fall Symposium: Emotional and Intelligent I, Orlando, FL.

Hudlicka, E. (2002). This time with feeling: Integrated Model of Trait and State Effects on Cognition and Behavior. *Applied Artificial Intelligence, 16*, 1-31.

Hudlicka, E. (2003). *Modeling Effects of Behavior Moderators on Performance: Evaluation of the MAMID Methodology and Architecture.* Paper presented at the BRIMS-12, Phoenix, AZ.

Hudlicka, E. (2004). Two Sides of Appraisal: Implementing Appraisal and Its Consequences within a Cognitive Architecture, *AAAI Spring Symposium: Architectures for Modeling Emotion* (Vol. TR SS-04-02). Stanford University, CA: AAAI Press.

Hudlicka, E. (2007). Reasons for Emotions. In W. Gray (Ed.), *Advances in Cognitive Models and Cognitive Architectures*. NY: Oxford.

Hudlicka, E. (2008a). Modeling the Mechanisms of Emotion Effects on Cognition, *AAAI Fall Symposium: Biologically Inspired Cognitive Architectures* (Vol. TR FS-08-04 pp. 82-86). Arlington, VA: Menlo Park, CA: AAAI Press.

Hudlicka, E. (2008b). What are we modeling when we model emotion? , *AAAI Spring Symposium: Emotion, Personality, and Social Behavior* (Vol. Technical Report SS-08-04, pp. 52-59). Stanford University, CA: Menlo Park, CA: AAAI Press.

Hudlicka, E. (forthcoming). *Affective Computing: Theory, Methods and Applications*. Boca Raton, FL: Taylor and Francis.

Izard, C. E. (1977). *Human Emotions*. NY: Plenum.

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

Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal & K. R. Scherer (Eds.), *The new handbook of Methods in Nonverbal Behavior Research* (pp. 65-135). NY: Oxford.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist 39*(2), 124–129.

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the Angry Decision Maker: How Appraisal Tendencies Shape Anger's Influence on Cognition. *Journal of Behavioral Decision Making, 19*, 115-137.

Leventhal, H., & Scherer, K. R. (1987). The relationship of emotion to cognition. *Cognition and Emotion, 1*, 3-28.

Lisetti, C., & Gmytrasiewicz, P. (2002). Can rational agents afford to be affectless? *Applied Artificial Intelligence, 16*(7-8), 577-609.

Loyall, A. B. (1997). *Believable Agents: Building Interactive Personalities.* CMU, Pittsburgh.

Mandler, G. (1984). *Mind and Body: The Psychology of Emotion and Stress*. New York: Norton.

Matthews, G. A., & Harley, T. A. (1993). Effects of Extraversion and Self-Report Arousal on Semantic Priming: A Connectionist Approach. *Journal of Personality and Social Psychology, 65*(4), 735-756.

Mehrabian, A., & (1995). Framework for a Comprehensive Description and Measurement of Emotional States. In *Genetic, Social, and General Psychology Monographs. Vol. 121.* (pp. 339-361).

O'Rorke, P., & Ortony, A. (1994). Explaining emotions. *Cognitive Science, 18*(2), 283-323.

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

Ortony, A., Norman, D., & Revelle, W. (2005). Affect and Proto-Affect in Effective Functioning In J. M. Fellous & M. A. Arbib (Eds.), *Who Needs Emotions?* NY: Oxford.

Ortony, A., & Turner, T. J. (1990). What's Basic About Basic Emotions? *Psychological Review, 97*(3), 315-331.

Panskepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. NY: Oxford University Press.

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

Plutchik, R. (1984). Emotions: A General Psychoevolutionary Theory. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion*. Hillsdale, N.J.: Erlbaum.

Prendinger, H., Descamps, S., & Ishizuka, M. (2002). Scripting Affective Communication with Life-like Characters in Web-based Interaction Systems. *Applied Artificial Intelligence, 16*(7-8), 519-553.

Prendinger, H., Saeyor, S., & Ishizuka, M. (2004). MPML and SCREAM: Scripting the Bodies and Minds of Life-Like Characters. In H. Prendinger & M. Ishizuka (Eds.), *Life-Like Charactes: Tools, Affective Functions, and Applications*. Berlin: Springer.

Reilly, W. S. N. (2006). *Modeling What Happens Between Emotional Antecedents and Emotional Consequents* Paper presented at the ACE 2006, Vienna, Austria.

Reilly, W. S. R. (1996). *Believable Social and Emotional Agents.* CMU, Pittsburgh.

Reisenzein, R. (1994). Pleasure-arousal theory and the Intensity of Emotions. *Journal of Personality and Social Psychology, 67*, 525–539.

Reisenzein, R. (2001). Appraisal processes conceptualized from a schema-theoretic perspective: Contributions to a process analysis of emotions. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*. NY: Oxford University Press.

Ritter, F. E., & Avramides, M. N. (2000). *Steps Towards Including Behavior Moderators in Human Performance Models in Synthetic Environments*: The Pennsylvania State University.

Ritter, F. E., Reifers, A. L., Klein, L. C., & Schoelles, M. J. (2007). Lessons from defining theories of stress for cognitive architectures. In W.Gray (Ed.), *Advances in Cognitive Models and Cognitive Architectures*. NY, NY: Oxford University Press.

Roseman, I. J. (2001). A Model of Appraisal in the Emotion System. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*. NY: Oxford.

Roseman, I. J., & Smith, C. A. (2001). Appraisal Theory: Overview, Assumptions, Varieties, Controversies. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*. NY: Oxford.

Rozin, P. (2003). Introduction: Evolutionary and Cultural Perspectives on Affect. In R. J. Davidson, K. R. Scherer & H.H.Goldsmith (Eds.), *Handbook of Affective Sciences*. NY: Oxford University Press.

Russell, J. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review, 110*(1), 145-172.

Russell, J., & Barrett, L. F. (1999). Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant. *Journal of Personality and Social Psychology, 76*(5), 805-819.

Russell, J., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research on Personality, 11*, 273-294.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures,* . Hillsdale, NJ: Lawrence Erlbaum.

Scherer, K., Schorr, A., & Johnstone, T. (2001). *Appraisal Processes in Emotion: Theory, Methods, Research*. NY: Oxford.

Scherer, K. R. (1992). Emotions are biologically and socially constituted: A response to Greenwood. *New Ideas in Psychology, 10*(1), 19-22.

Scherer, K. R. (2000). Emotions as Episodes of Subsystem Synchronization Driven by Nonlinear Appraisal Processes. In M. D. Lewis & I. Granic (Eds.), *Emotion, Development, and Self-Organization*. NY: Cambridge University Press.

Scherer, K. R. (2001). Appraisal Considered as a Process of Multievel Sequential Checking. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal Processes in Emotion: Theory, Methods, Research*. NY: Oxford.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? The information function of affective states. In K. Fiedler & J. P. Forgas (Eds.), *Affect, cognition,and social behavior* (pp. 44-62). Toronto: Hogrefe.

Sehaba, K., Sabouret, N., & Corruble, V. (2007). *An emotional model for synthetic characters with personality* Paper presented at the Affective Computing and Intelligent Interaction (ACII), Lisbon, Portugal.

Sloman, A., Chrisley, R., & Scheutz, M. (2005). The Architectural Basis of Affective States and Processes In J.-M. Fellous & M. A. Arbib (Eds.), *Who Needs Emotions?* . NY: Oxford University Press.

Smith, C. A., & Kirby, L. (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*. NY: Cambridge.

Smith, C. A., & Kirby, L. D. (2001). Toward Delivering on the Promise of Appraisal Theory. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal Processes in Emotion*. NY: Oxford.

Sonnemans, J., & Frijda, N. H. (1994). The Structure of Subjective Emotional Intensity. *Cognition and Emotion, 8*(4), 329-350.

Staller, A., & Petta, P. (1998). *Towards a tractable appraisal-based architecture for situated cognizers.* Paper presented at the 5th International Conference of the Society for Adaptive Behaviour (SAB'98): Grounding Emotions in Adaptive Systems Workshop., Zurich, Switzerland.

Velásquez, J. D. (1999). *An Emotion-Based Approach to Robotics* Paper presented at the IROS.

Appendix A: Glossary

A-1: Glossary of Terms Describing Affective States

Term	Definitions
Affective State	A transient mental state dominated by affect, emotion, or mood.
Affect (Note that some researchers use 'affect' as an umbrella term for all affective states – (e.g., Scherer; (Ortony et al., 2005); Rozin, 2003)	An undifferentiated mental state reflecting a positive or negative evaluation of current stimuli, reflecting valence only, and associated approach or avoidance behavioral tendencies. "pleasant and unpleasant feelings" (Frijda, 1994, p. 199) "perceived goodness or badness of something" (G. L. Clore, 1994) vs. "general umbrella term that subsumes a variety of phenomena such as emotion, stress, mood, interpersonal stance, and affective personality traits" (Juslin & Scherer, 2005) "energy in the mental system, [whose] aim was to be expressed" (Clore, 1994, p. 288) "the set of all valenced mental states, along with their associated physiological representations and behaviors" (Rozin, 2003)
Emotion	A transient mental state lasting seconds to minutes, typically of moderate to high intensity, involving an evaluation of the current situation with respect to active goals, and associated expressive and behavioral manifestations. "state of affective appraisal of some object (external or internal), linked to a change in action readiness bearing on that object" or "involve a change in action readiness and are about something" (Frijda, 1994, p. 199) "a sequence of state changes in each of five – functionally defined – organismic subsystems (cognitive, autonomic nervous system, motor system, motivational system, monitor system) occurring in an interdependent and interrelated fashion in response to the evaluation of a stimulus, an event, or intraorganismic changes of central importance to the major needs and goals of the organism" (K.R Scherer, 2000) p. 74). "interpreted feelings" (Ortony, Norman and Revelle, 2005, p. 174)
Basic Emotions	Emotions involved in basic survival within the environment. The exact set varies to some extent, but typically includes: fear, anger, joy, sadness, disgust, and surprise. "Basic emotions are defined as corresponding to inborn, phylogenetically selected, neuro-motor programs. They are in limited number and are universal reactions (universality operates across ages, across cultures and across species)." (Panskepp, 1998, p. 46) "each of the basic emotions is produced by an innate hardwired neuromotor program with characteristic neurophysiological, expressive, and subjective components" (Ellsworth & Scherer, 2003, p. 574)
Non-Basic or Complex Emotions	Those emotions not considered basic, having a larger cognitive component and triggering and manifestation variabilities, and including the social emotions (see below). Examples of other non-basic emotions include love, awe, contentment and schadenfreude.
Social Emotions	Emotions involved in coordinating interpersonal behavior and typically requiring a representation of the self within its social milieu – include pride, shame, guilt, jealousy, envy, embarrassment, contempt and jealousy
Moods	A transient mental state lasting hours to days to months, involving a generalized, feeling state, typically of lower intensity than emotion, and often associated with one of the emotions; e.g., happy, sad, angry, fearful. The specific trigger is often not apparent and the behavioral reactions are typically diffuse: not directed at any particular individual or object, but rather at the 'world at large'.

	"diffuse states of feeling or action readiness or both" (N. H. Frijda, 1994), p. 199)
Feelings	Non-specific subjective mental states involving a conscious awareness of a combination of physical, mental and / or emotional characteristics. Feelings can be affective (good, bad, jealous), physical (hungry, tired), cognitive (confused, certain) (certain, confused). "readouts of the brain's registration of bodily conditions and changes—muscle tension, autonomic system activity, internal musculature (e.g., the gut), as well as altered states of awareness and attentiveness" (Ortony, Norman & Revelle, 2005, p. 174).
Sentiments	"long-lasting emotional states of relating to other people" (e.g., love, trust or distrust, arrogance or deference) (Oatley, 2004, p.4) (Similar to Scherer's "interpersonal stance" (Scherer, 2000).

Table A-2: Glossary of Terms Used in Emotion Generation and Cognitive Appraisal

Terms / Synonyms	Definition
Appraisal, Construal	Interpretation of an event, situation, or object in light of the agent's concerns, goals, behavioral norms and preferences.
Valenced reaction	An affective reaction resulting from a appraisal / construal.
Emotion elicitors, emotion triggers, emotion antecedents	Situation, event or object that triggers an emotion
Emotion consequents, emotion effects	Effects of emotions, both internal (perception, cognition) and external (expressions, specific action)
Concerns	"motivating dispositions" (Frijda, 1986) of an agent; its goals, needs, likes/dislikes
Goal – generic	"a desired state of affairs that, should it obtain, would be assessed as somehow beneficial to the agent" (Elliot, 1992)
Goal – OCC theory	Agent's concerns that are impacted by events (vs. acts by other agents or attributes of objects)
Drive	Low-level concern, reflecting basic biological needs such as hunger, thirst, rest
Components (of emotion)	In componential theories of emotion (e.g., Scherer), the individual modalities comprising an emotion; in cognitive appraisal, components can refer to the individual appraisal variables
Core affect	"Neurophysiological state consciously accessible…blend of hedonic (valence) and arousal values" (Russell, 2003)
Valence, Pleasure	Measure of the degree of desirability (of a stimulus) (positive vs. negative)
Arousal	Measure of the degree of excitation in an agent, generally reflects autonomic nervous system activity, and speed and intensity of processing
Dominance (vs. Submissiveness)	Attitude of agent toward the world and other agents.
PAD	Pleasure, Arousal, Valence (used in dimensional theories of emotion to characterize the emotion space)
Local variables, emotion eliciting conditions	In OCC theory: evaluative criteria used by appraisal processes (e.g., desirability, praiseworthiness, attractiveness). Different emotions associated with different local variables (analogous to appraisal variables)
Global variables	In OCC theory: evaluative variables applicable to all emotions and influencing intensity, includes physiological arousal.
Appraisal variables, appraisal dimensions	In componential and cognitive appraisal theories: Domain-independent evaluation criteria or features, used to evaluate a stimulus and its relevance to the agent's well-being (analogous to OCC local variables.
Stimulus evaluation checks, appraisal detectors	Processes calculating the values of appraisal variables: stimulus evaluation checks (Scherer); appraisal detectors (Smith & Kirby, 2000)
Appraisal state	A particular configuration of appraisal variable values, corresponding to the agent's evaluation of the current set of stimuli.
Object domain	Domain within which the agent exists, and where emotion-eliciting triggers occur; by definition domain-specific (Elliot, 1992).
Mental domain	Internal, mental representations necessary to mediate appraisal, such as goals and expectations.
Abstract domain	Domain-independent evaluative dimensions (local variables in OCC, appraisal variables in componential theories) (Elliot, 1992).
Affective domain	The end results of the emotion generation process, consisting of the derived affective states (emotions, moods).

Table A-3: Example of a Frame Representing an Emotion Instance

Attribute	Content	Example
Affective State type	{affect, emotion, mood, attitude}	Emotion
Valence	Some value between –n and +n, where n is typically 1	1
Emotion type	Name of emotion	Joy
Intensity / Activation Level	Some value between 0 and n, where n is typically 1	.5
Underlying dimensions, if dim model used (PA or PAD)	{Arousal, Valence, Dominance}, represented by values between –n and +n, where n is typically 1	(.5,-1,0)
Underlying appraisal variables, if componential model used	List of specific appraisal variables and their values	Novelty = 1 Goal congruence = high Etc.
Time when first created	T where t > 0	3
Current time	$T_{current}$, where $T_{current}$ >= T	
Duration / Decay function	Specific decay function for this emotion type	2 minutes / Exponential
Eliciting Triggers (may be further categorized into types, as per OCC theory)	Pointers to structures containing list of triggers (e.g., events)	e.g., Event_12; Situation_5
Affected goals/concerns (internal evaluative criteria) (may be further categorized into types, as per OCC theory; e.g., goals, standards, preferences)	Pointers to structures containing list of goals/concerns	e.g., Goal_22; Behavioral_Norm_42
Direction / Target of emotion-triggered behavior	List of agents (including self) and objects	Agent_007