What Are We Modeling When We Model Emotion?

Conference Paper · January 2008

Source: DBLP

1 author:

Eva Hudlicka
Visiting Lecturer, University of Massachusetts-Amherst

What Are We Modeling When We Model Emotion?

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. Architectures are being built both to elucidate 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'. 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 the theoretical and data requirements, and the representational and reasoning alternatives, in emotion modeling. Identification of a set of generic computational tasks is also a good starting point for a systematic comparison of alternative approaches.

Introduction and Objectives

The past 15 years have witnessed a rapid growth in computational models of emotion and affective 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'. '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 factors are modeled. 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 factors. 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 can thus be contrasted with other terms describing affective phenomena: *moods*, sharing many features with emotions but lasting longer (hours to months); *affective states*, undifferentiated positive or negative 'feelings' and associated 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)). Some models also represent permanent affective personality *traits* (e.g., extraversion, neuroticism), or a variety of 'mixed' mental states that involve both cognitive and affective components (e.g., attitudes).

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.

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?

A 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? The building blocks can thus serve as basis for what Sloman calls 'architecture based definition of emotion" (Sloman et al., 2005).

A note on terminology: I use 'agent' to mean any type of autonomous entity, whether biological or synthetic; I use 'EMOTIONS' to mean the broad category of transient states that includes the various evaluative states identified by emotion researchers, including undifferentiated affective states (positive / negative), emotions proper, and moods.

What Are EMOTIONS?

Definitions 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, emotion researchers do agree on a high-level definition of EMOTIONS, as the "evaluative judgments of the environment, the self and other social agents, in light of the agent's goals and beliefs".

EMOTIONS as Multi-Modal Phenomena 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 modality (manifesting emotions in agents), it is important to recognize that both the physiological, and the experiential modalities, also play critical roles (Izard, 1993).

Modeling Emotion Generation

Cognitive Appraisal Theories 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, 2005; 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".

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. 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 (e.g., Arnold, 1960; Frijda, 1986; Lazarus, 1984; Mandler, 1984; Roseman & Smith, 2001; Scherer et al., 2001; Smith & Kirby, 2001). The most influential appraisal theories in computational modeling are those that are cast in 'computation-friendly' terms. The first of these was a theory proposed by Ortony and colleagues, now referred to as the OCC model (Ortony et al., 1988). More recently, appraisal models proposed by Scherer, and Smith and colleagues, have become the basis for computational appraisal models (Scherer et al., 2001; Smith & Kirby, 2000).

OCC remains the most frequently implemented model of appraisal (e.g., Andre et al., 2000; Bates et al., 1992; Reilly, 2006). It provides a rich taxonomy of triggers and resulting emotions, focusing on fundamental distinctions among three types of triggers, and corresponding types of emotions: *events* (event-based emotions such as desirability, hope, fear), *acts by other agents* (attribution emotions such as anger), and *characteristics of objects* (attraction emotions such as like, dislike). An OCC-based appraisal model proceeds through a sequence of steps as it classifies a trigger within this taxonomy, eventually generating a specific emotion.

More recently, appraisal theories of Scherer, and Smith and Kirby, have been used as the basis of computational models. These theories emphasize a set of explicit, domain-independent features of the stimuli, the *appraisal dimensions*, which include *novelty, valence, goal relevance and goal congruence, responsible agent, coping potential, and norms and values.* (Similar variables are also identified in the OCC theory, but structured somewhat differently within the OCC emotion taxonomy.) These models first extract the appraisal dimension values from the triggering stimuli, generating a point within the space defined by these dimensions, which is then mapped to a particular emotion. Theories of Smith and colleagues are similar to Scherer's, but emphasize the role of the agent's assessment of its ability to successfully handle the situation (coping). 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 architectures in general (Ortony et al., 2005; Sloman et al., 2005)).

From Theories to Models Ideally, appraisal theories would provide sufficient details for a computational model. This means answering questions such as:

(1) 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 a series of intervening appraisal dimensions (e.g. novelty, valence, responsible agent).
(2) How are external stimuli integrated with internal stimuli (recalled or anticipated events and situations) in triggering emotions?
(3) What are the distinct stages of the appraisal process and the functions implemented in each stage? Is there any variability in these stages, as a function of the specific emotion, individual, or context?
(4) What are the dependencies and interactions among the distinct processes implementing appraisal?
(5) What factors influence emotion intensity, and what is the nature of this influence?
(6) What are the 'emotion dynamics' - the ramp-up and decay rates of individual emotions? How do these vary by individuals, emotions and contexts?
(7) Can multiple emotions be generated, and, if not, how should potentially conflicting triggers be integrated into a single emotion?
(8) What cognitive structures are necessary to support appraisal, and what should the nature and complexity of these structures be (e.g. goals, expectations, plans).
(9) What levels of resolution are necessary for a particular modeling application (e.g., sensorimotor, schematic, or conceptual)?

Unfortunately, existing theories do not provide answers to all of these questions. In fact, it is frequently the act of model construction itself that motivates more refinements of the associated psychological theories. Nevertheless, the questions above do provide a basis for defining a set of computational tasks necessary to implement emotion generation via appraisal. These are described below, along with examples from existing models, illustrating alternative approaches.

Stimulus-to-Emotion Mapping The primary task is to map the triggering stimuli (emotion antecedents) onto the resulting emotion (or mood, or affective state), which reflects the agent's evaluation of these stimuli, in light of its goals and beliefs. Existing empirical data provide a rich source for these mappings; e.g., we know that possibility of bodily harm triggers fear; obtaining a desired object causes happiness, etc. Complexities arise, however, as we begin to consider factors such as: (1) individual and cultural variability, and the influences of agent's current emotional state or situational context; (2) stimuli that reflect not only the agent's immediate external environment, but also recalled and imagined events and situations, which may require complex representations of the self and other agents; (3) effects of dynamic cognitive constructs, such as goals, expectations, beliefs (both long-term knowledge of the world and current assessments of existing situations); (4) and the influences of the agent's current emotions, moods and personality traits. As we begin to refine the model requirements to accommodate these factors, we quickly find that the supporting theories and data may not be available.

This mapping task can be implemented with simple black-box models, that map the triggering stimuli onto the resulting emotion(s); either directly ("growling dog causes fear"), or via a set of intervening domain-independent appraisal variables (i.e., {high novelty, negative valence, negative goal congruence, low coping potential} cause fear). Both approaches are being used to model appraisal, with a recent trend being towards the use of the domain-independent appraisal dimensions, which enables a higher resolution of the emotion space, and affords a degree of domain independence. It should be emphasized that while the {appraisal dimension}-to-{emotion} mapping may be relatively simple, extracting the values of the appraisal dimensions from the domain data is far from trivial, and may require complex AI representational formalisms and reasoning (e.g., determining whether a particular situation in the environment is conducive to the agent's goal may require both complex situation assessment, and the management of intricate planning, causal and goal structures).

The mapping task can also be implemented using more process-based models, which attempt to explicitly represent the underlying mechanisms mediating the appraisal process. Benefits of process models include the potential for elucidation and refinement of the underlying mechanisms, and increased robustness.

A number of additional design choices must be made, regarding the nature of the internal, dynamic constructs required, complexity of causal and planning structures, necessity for an explicit representation of time and 'self', etc. For example, fear and hope are "prospect-based" emotions, requiring representations of future states. Regret and remorse require representations of histories of the self, and / or others and the world, and the ability to perform counterfactual reasoning. All of these have implications for the representational requirements and the selection of formalisms.

Emotion Dynamics: Emotion Intensity and Ramp-Up and Decay Rates The seemingly simple attribute of *emotion intensity* also reflects a high degree of complexity. Not only must we define the fundamental formulae for calculating intensity, based on the types, and characteristics, of the triggering stimuli, but we must then integrate the 'current' intensities with those of existing emotions and moods, to ensure smooth and appropriate transitions among states. This must take into account possible differences in the decay rates of different emotions, which are subject to a variety of influences that have not yet been identified or quantified to the degree required for computational modeling. Reilly (2006) discusses some alternatives for modeling emotion dynamics.

Most existing models of appraisal use relatively simple formulae for calculating emotion intensity, typically focusing on desirability and likelihood; e.g., [desirability * likelihood] (Gratch & Marsella, 2004), [desirability * (change in) likelihood] (Reilly, 2006). A number of complexities are typically not addressed. For example, Reilly (2006) points out the need for representing asymmetry of success vs. failure; in other words, for different types of individuals (and different goals) success may be less (or more) important than failure; e.g., extraversion is associated with reward-seeking whereas neuroticism is associated with punishment-avoidance. Modeling of these phenomena requires distinct variables for success (desirability of an event, situation or world state) vs. failure (undesirability of the same).

Directly related to the intensity calculation is the calculation of the emotion ramp-up and decay rates, which brings up a question regarding the extent to which emotions represent self-sustaining processes, that must 'run their course'. Reilly summarized current approaches to decay calculation as being linear, logarithmic, exponential, or "some arbitrary monotonically decreasing function over time" (Reilly, 2006).

Unfortunately for modelers, emotion dynamics are not well understood, and the data for precise calculations of intensities and ramp-up and decay rates are not available. Existing empirical studies provide qualitative data at best. Variability of these processes across emotions and individuals, while documented, has also not been quantified; e.g., high neuroticism rate predisposes individuals towards faster and more intense negative emotions; anger appears to decay more slowly than other emotions (Lerner & Tiedens, 2006). Even more importantly, some researchers point out that the appraisal dimensions identified for emotion differentiation may not be the same as those that "allow prediction of duration and intensity", and that "the current set of appraisal dimensions may be incomplete" (Scherer, 2001, p. 375).

Combining Multiple Emotions Emotions do not occur in isolation. Multiple emotions may be generated by the appraisal process, and existing emotion(s) must be combined with newly-generated emotion(s). At their maximum intensity, we may feel, and express, a single emotion. However, more typically, multiple emotions interact to form the subjective 'feeling' experience and to influence cognitive processing and behavior selection. These phenomena are not well understood, let alone quantified, to the degree required for modeling.

Reilly has analyzed several existing approaches to combining similar emotions 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.

No analogous analysis exists for combining opposing emotions. Nor do existing theories and empirical data provide much help. 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 right 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? 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.

Modeling Emotion Effects

It is useful to divide emotion effects into two categories: the visible, often dramatic, behavioral expressions, and the less visible, but no less dramatic, effects on attention, perception and cognition. Majority of emotion models focus on the former.

While technically challenging, the behavioral effects are easier from a modeling perspective, due to the large body of empirical data regarding the visible manifestations of particular emotions. We know, in general, how various basic emotions are expressed in terms of visible behavior: facial expressions, quality of movement and gestures, behavioral choices. (As with emotion generation, the degree of variability and complexity increases as we move from the fundamental emotions such as fear, joy, anger, to the more cognitively-complex emotions such as pride, shame, jealousy etc.).

Much less well understood, thus less frequently modeled, are the internal effects emotions exert on the perceptual and cognitive processes that mediate adaptive, intelligent behavior; both the fundamental processes (attention, working memory, long-term memory recall and encoding), and higher-level processes such as situation assessment, counterfactual reasoning, problem-solving, goal management, decision-making, learning, and action selection. The focus below is on modeling these 'internal' emotion effects.

The effects of emotions on these processes have been studied by psychologists, and some data are available; e.g., positive emotions induce more global thinking and use of heuristics; anxiety reduces attentional and working memory capacities, biases attention towards detection of threatening stimuli, and biases interpretive processes towards higher threat assessments; mood induces recall biases; negative affect reduces estimates of control, and induces more analytical thinking (Isen, 1993; Mineka et al., 2003). However, theories of the mechanisms of these influences are not nearly as well elaborated as those for emotion generation via cognitive appraisal. The translation from theory to computational tasks, outlined above for the appraisal models, is thus more challenging for models of emotion effects on cognition.

Theories of Emotion Effects The few existing theories postulating specific mechanisms of emotion effects on cognition can be classified into one of two broad categories: *spreading activation models*, and *parameter-based models*, which suggest that affective factors act as parameters inducing variabilities in cognitive processes (and behavior).

Spreading activation has been proposed to explain several phenomena in emotion-cognition interaction, particularly *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 (Hudlicka, 2003; 2007; Ritter et al., 2007; Sehaba et al., 2007; Belavkin & Ritter, 2004;).

These 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).

From Theories to Models As with appraisal theories, theories of emotion effects would ideally provide sufficient detail to construct computational models. Questions we would like answered include:

(1) 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?

(2) How are contents and organization of long-term memory structures affected?
(3) How is cognitive appraisal affected by emotions?
(4) 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?
(5) 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).
(6) How and when are the influences of multiple emotions, moods and traits combined?
(7) Are there distinct types of processes that mediate these influences? What are the interactions and dependencies among these processes?
(8) Last, but not least: Can we obtain sufficient data about these internal processes and structures to enable construction of computational models?

Unfortunately, the theories attempting to explain the mechanisms of emotion effects on cognition are less-well elaborated than theories of cognitive appraisal, and do not provide adequately detailed answers to these questions. Nevertheless, these questions do provide a basis for defining a candidate set of computational tasks necessary to implement emotion effects on attention, perception and cognition. These are described below, along with some examples from existing models.

Emotion-to-Cognitive-Processes Mappings A series of tasks needs to be defined here for each of the documented effects of emotion on the distinct cognitive structures and processes. We can begin by identifying tasks that focus on the fundamental processes underlying high-level cognition: effects on attention (speed, capacity and accuracy), working memory (encoding and recall speed, capacity and accuracy), and long-term memory (encoding, recall, content, organization). Additional specific tasks depend on the objective and structure of a particular model; e.g., a see-think-do model will require identifying the computational tasks necessary to implement emotion effects on situation assessment, goal selection, action selection, and execution monitoring.

Existing data provide some help in defining these tasks, at least in qualitative terms, and a number of models implement some of these. For example, Breazeal's Kismet uses emotion to focus attention, prioritize goals, and select action, within a broader robot architecture that also includes emotion generation and expressive behaviors (Breazeal, 2005). Hudlicka's MAMID cognitive-affective architecture encodes emotions in terms of parameters, which then induce changes in capacity, speed, and biases in a number of cognitive processes, including attention, working and long-term memory, situation assessment, goal management and action selection, as well as cognitive appraisal processes themselves (Hudlicka, 2003; 2007). A similar parameter-based approach, implemented within ACT-R, has been used to model action selection (via conflict resolution) (Belavkin & Ritter, 2004), and stress effects (Ritter et al., 2007). Several models of emotion effects on behavior selection use a decision-theoretic formalism, where emotions bias the utilities and weights assigned to different behavioral alternatives (Busemeyer et al., 2007; Lisetti & Gmytrasiewicz, 2002).

Determining Emotion Effect Magnitude Going beyond the qualitative relationships typically identified in empirical studies (e.g., anxiety biases attention towards threatening stimuli) is more difficult, since existing empirical data do not provide sufficient information for calculating the exact magnitudes of the observed effects. More accurate data are available at the periphery of the cognitive system (attention and motor control tasks), and for simple laboratory tasks. In the majority of existing models, quantification of the available qualitative data is therefore more or less ad hoc, typically involving some type of linear combinations of the weighted factors, and requiring significant fine-tuning to adjust model performance; e.g., in MAMID, the affective factor influences on module capacity are represented by weighted terms that contribute positively or negatively to the overall module capacity, depending on the existing empirical findings (e.g., anxiety reduces capacity) (Hudlicka, 2008). Current trend to combine empirical studies with computational modeling efforts is promising, and will hopefully provide some of the necessary quantitative data.

Integration of Multiple Emotions As is the case with effects magnitude, existing empirical studies generally do not provide information about how to combine multiple effects, or how these may interact. This requires that the modeler combine known qualitative data in a somewhat ad hoc manner, and tune the resulting models to obtain the desired behavior. As was the case with appraisal, a number of issues must be addressed in combining similar, different or opposing effects; both regarding the stage of processing where these effects should be integrated, and regarding the exact manner of this integration.

For both of the tasks above, data for the internal processes and structures (e.g., effects on goal prioritization, expectation generation, planning) are more difficult to obtain and quantify, due to the lack of direct access and assessment, and the transient nature of emotions and the affected cognitive constructs. This may indeed provide a limiting factor for models of these phenomena. Currently, the degree of resolution possible within a computational model far exceeds the degree of resolution of the data we are able to obtain about these processes, resulting in models that are highly underconstrained, and thus limited in their explanatory capabilities.

Related Work

A number of researchers have addressed the issue of systematizing emotion modeling, both at the individual task level, and at the architecture level. Reilly's work is most closely related to my attempt to identify individual computational tasks required for emotion modeling, and identifies many of the tasks identified for modeling cognitive appraisal (2006). Lisseti and Gmytrasciezicz identified a number of high-level components of emotion required for computational models, in their Affective Knowledge Representation scheme (Lisetti & Gmytrasiewicz, 2002). Canamero discusses design requirements for affective agents, focusing on the role of emotion in action selection (2002). In terms of architectures, Sloman and colleagues have discussed the architectural requirements for adaptive behavior in general, which includes emotion, and implemented a computational model: CogAff model (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 et al., 2005).

Summary and Conclusions

Recognizing the lack of consistent terminology and design guidelines in emotion modeling, this paper proposes an 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 involve, for both processes: defining a set of *mappings* (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 in the triggering stage (antecedents), and in the emotion effects stage (consequences).

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 representative examples were discussed. All three of these limitations will be partially addressed in two forthcoming publications (Hudlicka, 2008; 2009).

It is hoped that the analysis presented here will stimulate more focused dialogue, a refinement of the proposed analytical framework (especially by including the additional affective modalities), and contribute to a definition of a catalogue of available tools and methods, which will eventually lead to more systematic design guidelines for modeling these complex phenomena, an form a basis for their evaluation and validation.

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. In Proceedings of IWAI, Siena, Italy.

Arnold, M. B. (1960). *Emotion and personality.* NY: 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.

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

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

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 and M. A. Arbib (Eds.), *Who Needs Emotions?* NY: Oxford.

Busemeyer, J. R., Dimperio, E., & Jessup, R. K. (2007). Integrating emotional processes into decision-making models. In W.Gray (Ed.), *Integrated Models of Cognitive Systems*. NY: Oxford.

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

Fellous, J. M. (2004). From Human Emotions to Robot Emotions. In Proceedings of the AAAI Spring Symposium 2004-Architectures for Modeling Emotion, Stanford University, Palo Alto, CA.

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

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. In Proceedings of the AAAI Fall Symposium: Emotional and Intelligent I, Orlando, FL.

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

Hudlicka, E. (2004). Two Sides of Appraisal: Implementing Appraisal and Its Consequences within a Cognitive Architecture. In Proceedings of the AAAI Spring Symposium: Architectures for Modeling Emotion, Stanford University, Palo Alto, CA.

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

Hudlicka, E. (2008). Guidelines for Modeling Affect in Cognitive Architectures *Submitted for publication to Journal of Cognitive Systems Research (Also: Report # 0802, Psychometrix Associates, Inc. Blacksburg, VA).*

Hudlicka, E. (2009). *Affective Computing: Theory, Methods and Applications*. Boca Raton, FL: Taylor & Francis: CRC Press. Forthcoming

Isen, A. M. (1993). Positive Affect and Decision Making In J. M. Haviland and M. Lewis (Eds.), *Handbook of Emotions.* NY: The Guilford Press.

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

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.

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

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.

Mineka, S., Rafael, E., & Yovel, I. (2003). Cognitive Biases in Emotional Disorders: Information Processing and Social-Cognitive Perspectives. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Science*. NY: Oxford.

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 and M. A. Arbib (Eds.), *Who Needs Emotions?* NY: Oxford.

Reilly, W. S. N. (2006). Modeling What Happens Between Emotional Antecedents and Emotional Consequents In Proceedings of ACE 2006, Vienna, Austria.

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: Oxford.

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

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

Sehaba, K., Sabouret, N., & Corruble, V. (2007). An emotional model for synthetic characters with personality. In Proceedings of 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 and M. A. Arbib (Eds.), *Who Needs Emotions?* NY: Oxford.

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 A. S. K.R. Scherer, T. Johnstone (Ed.), *Appraisal Processes in Emotion*., NY: Oxford.

Velásquez, J. D. (1999). An Emotion-Based Approach to Robotics In *Proceedings of IROS*.