Advances in Cognitive Systems 2 (2012) 59-76 Submitted 9/2012; published 12/2012

Computational Models of Emotion and Cognition

Jerry Lin JERRYLIN@USC.EDU
Marc Spraragen SPRARAGE@USC.EDU
Michael Zyda ZYDA@USC.EDU
Computer Science Department, University of Southern California, Los Angeles, CA 90089 USA

Abstract

In this paper, we seek to review the broad landscape of research in computational emotions and cognition. We begin by classifying and organizing an enumeration of recent models and systems and then discuss some of the landmark models from the literature, such as EMA and WASABI. We then discuss open problems with the current state of research. These issues are standardizing criteria for evaluation of models, the complexity and breadth of the domain, and the need to implement a working system which addresses integration with more of the rich history of AI research. We also provide suggestions for future research, particularly standardization to facilitate community collaboration.

1. Introduction and Background

Emotion and cognition have long been thought to have important interaction, from Plato's chariot allegory, characterizing reason and emotion as horses pulling a chariot and the charioteer having to direct each horse to move together in order to reach enlightenment. More recently, arguments have been made that thought is essential for emotion (Lazarus, 1982), and emotion is essential for thought (Damasio, 1994). Overall, a large amount of research within the last few decades has yielded ideas and data about the nature of this relationship. For instance, appraisal theories form a framework for emotion generation, (Ortony, Clore, & Collins, 1988; Schorr, 2001), and mood-congruent recall and decision theories (Bower, 1983) capture some of the cognitive effects of emotion. However, there are still open questions and difficulties, providing fertile ground for further computational models for experimentation and detailed understanding.

There is much confusion regarding emotion terminology, and this is in part due to emotion being a common part of everyday life leading to intuitive meanings for terms which typically vary from technical definitions. Furthermore, there is no consensus on the technical definitions themselves, leading to definition sections as seen here. For this paper, we will define six terms as follows:

Affect Any information (emotion, feeling, mood) used to inform one or more cognitive processes

Appraisal The process of making judgments (appraisals) about the relationship between perceived events and one's beliefs, desires, and intentions (Lazarus, 1991)

Cognition Mental processes that have to do with the acquisition, alteration, and comprehension of knowledge; such as recall, inference, learning, and planning

Emotion Cognitive data arising from events (internal and external) used to inform responses, and attributed to concepts and states

Feeling The subjective experience of an emotion or set of emotions

Mood An overall state of emotion which is sustained over longer periods of time and is less changeable than emotions themselves

Computational models of emotion and cognition which we address are those that try to explain emotion in the context of its intimate relationship with cognition. These models are distinguished from those in psychology and cognitive science by having a level of detail about the processes and data involved to be implemented on a modern computer.

In the 19th century, the psychologist William James and others theorized that emotions were brought on by physiological reactions to situations. James's theory was a precursor to appraisal theory (Scherer, 1999), whose proponents also view emotions as effects of reactions to situations, though with less of a focus on physiological reactions. Appraisal theory tends to dominate among computational models of emotion due to its emphasis on emotions as computable artifacts. It also is the center of modern theories of emotion as it frames emotions arising from cognition and thus explaining the intimate relationship between the two. There are still many major holes regarding the interplay between emotion and cognition, however, which computational models of emotion can help to elucidate.

Appraisal theory is dominant in the community of computational emotional modeling, although other schools of thought have also made an impact in that arena. The theory was developed as a means to predict individual human emotions given particular situations (Arnold, 1960; Lazarus, 1966; Scherer, 1999). The basis of the theory is that a person can appraise (i.e., evaluate) an entity, concept, event, or situation with respect to the appraiser's beliefs, desires, and intentions. The dimensions along which these appraisals are made are called appraisal variables and a certain combination of appraisal variable values predictably gives rise to a distinct emotion. The mapping from appraisal variables to emotion has been termed "affect derivation" (often coupled with or subsuming a derivation of emotional intensity). For instance, a shark encounter might be appraised by a swimmer as likely to result in serious physical harm, and this appraisal would generate an intense emotion of fear in the person.

One way in which appraisal theories differ from one another is in the number, breakdown, and definition of appraisal variables accounted for by each theory, and the ways in which the variables combine to predictably generate labeled emotions. However, most appraisal theories share some fundamental concepts: valence (positive/negative rating of event, object, or situation) is present or inferable among many theories' appraisal variables (Ortony, Clore, & Collins, 1988; Blascovich & Mendes, 2000; Scherer, 1999) and arousal (intensity of feeling) is measured as an appraisal variable in several theories or else assumed to be a factor in generating emotional response upon appraisal of a situation relevant to the agent and its goals (Marsella, Gratch, & Petta, 2010).

The concept of coping potential (ability to deal with a situation either by action or cognition) is also common though in various forms. Coping itself may include taking direct action regarding the situation, or cognitive redefinition of one's beliefs, desires, or intentions; for example, the "sour grapes" approach of reappraising a negative situation as a positive. Lazarus (1966) developed the similar concept of "primary" vs. "secondary" appraisals: primary appraisal takes in a situation's significance, and secondary appraisal (coping potential) assesses the ability to deal with the situation. Frijda (1987) relates emotions to "action tendencies," with emotional cues providing constraints on the next set of decisions or actions made by an agent. For instance, fear may limit action tendencies to adversive behavior.

The Ortony, Clore, and Collins (1988) appraisal theory, often referred to as OCC theory, categorizes emotions based on appraisal of pleasure / displeasure (valence) and intensity (arousal). To more specifically predict emotion generation, it breaks down valence appraisal into three categories based on what is being appraised: desirability (of an event), praiseworthiness (of an action), and like/dislike (of an entity). Also, actions and events may be further differentiated by an "attribution" variable: for instance, was an action taken by (or did an event affect) oneself, or another? Appraisal across these variables defines different specific emotions; for instance, a positive-valenced appraisal of an action attributable to oneself might create an emotion of pride in the appraiser, whereas the swimmer's appraisal of the shark encounter as above would be as a negative event attributable to the shark, with prospective negative consequences for the swimmer, thus producing fear of the shark.

Several researchers have devised high-variable-count appraisal theories that map specific configurations and values of appraisal variables to a range of generated emotions. One such map (Scherer, 2001) is summarized in Table 1.

Table 1. Appraisal dimensions from Scherer's appraisal theory.

	Sequence	Joy	Fear	Anger	Sadness	Disgust	Shame	Guilt
Expectedness	1	Open	Low	Open	Open	Open	Open	Open
Unpleasantness	1	Low	High	Open	Open	V.High	Open	Open
Goal Hindrance	1	V.Low	High	High	High	Open	Open	Low
External Causation	2	Open	Ext.	Ext.	Open	Ext.	Int.	Int.
Coping Potential	3	Medium	V.Low	High	Low	Open	Open	Open
Immorality	4	Open	Open	High	Open	Open	Open	V.High
Self-Consistency	4	Open	Open	Low	Open	Open	V.Low	V.Low

Other theories outline a distinction between non-cognitive and cognitive appraisal. One difference is that cognitive appraisal processes (like inferring the cause of a situation) are generally slower than appraisals based on direct sensory feedback (like physical pain). Bechara et al. (2000) expresses these different appraisal levels as somatic (primary) and recalled (secondary) "emotion inducers". Some researchers (Becker-Asano & Wachsmuth, 2009) use this distinction to define certain emotions as "secondary" – only able to arise following some cognitive processing. For instance, anger at a person most likely stems from cognitively attributing an action or event (previously appraised as unpleasant) to that person, such as a "shark attack" revealed to be another swimmer playing a practical joke.

Part of the difference between theories that postulate two levels of appraisal and theories which only identify one (fast) appraisal level is semantic in nature: the two-level theories include deliberative cognitive processing (e.g., inference or recall) as part of "secondary appraisal," whereas the one-level theories limit appraisal to quick situational evaluation of sensed situations and cognized situations alike. The one-level theories view the cognitive processing of events as coping, instead of as appraisal (Gratch & Marsella, 2004). Also, according to Leventhal and Scherer (1987), there is not a clear line between calling a given appraisal process cognitive or non-cognitive. The line can be further blurred in that routine appraisals of a particular situation can enable future appraisals (recognition response) of that situation to be quicker and more reflexive than an initial, purely cognitively based appraisal (Lehrer, 2009).

Dimensional theories of emotion generation are similar to appraisal theories in that both map emotion-evoking events to emotional states. The main difference is that while appraisal theories relate discrete appraisal elements to discrete emotional states, dimensional theories posit a non-relational "core affect" (or mood) state tracked as a single uniquely determined point along a number of continuous, orthogonal dimensions (Marsella & Gratch, 2009). The two-dimensional circumflex theory (Russell, 1980) places various emotional states around the axes of pleasure and arousal. The PAD dimensional theory (Mehrabian & Russell, 1974) is named for its three dimensions of Pleasure (valence), Arousal, and Dominance (defined as the degree to which a person feels powerful or in control of the situation, analogous to coping potential in appraisal theory). PAD is analogous to a 3-dimensional expansion of the circumflex. For instance, in PAD theory both anger and anxiety arise from similar low-Pleasure and high-Arousal events. However, anger and anxiety are on opposite sides of the Dominance dimension: an anxious person feels less in control of their situation than does an angry person. Like Challenge and Threat theory, but unlike the concept of primary vs. secondary appraisals, PAD posits both Pleasure and assessment of coping potential (dominance) at the same level.

Cognition has been an area of intense study in the artificial intelligence community for a relatively long time now. Important work in cognitive architecture (Langley, Laird, & Rogers, 2009; Laird, 2008; Anderson et al., 2004) has studied cognition in the integrated context of an intelligent system. Emotion is believed to have developed in the reptilian brain before higher levels of cognition (Lazarus, 1982), and has a natural place informing the "higher" levels of the architecture.

2. Representative Models and their Properties

The impact of recent emotion-related human psychological and cognitive studies has contributed to an increase in computational modeling of emotion and cognition, allowing subcategories of AI systems to form. Affect-antecedent systems, for example, how thinking of a plan changes affective state (Gratch, 1996), or focus on basic expressions to register the presence of emotions. Affect-consequent systems–computational models of the effects of emotion on cognition (Gratch, Marsella, & Petta, 2009)–may be categorized along several criteria. One clear demarcation is between "behavior-consequent" and "cognitive-consequent" models, although many systems include both of these functions. A behavior-consequent model maps an agent's emotional state to embodied physical actions or other direct outward or social expression, for instance smiling when happy or turning on a light if afraid of the dark. Behavioral-consequent models are often used to synthesize human-like emotional or social behavior in embodied robots like Kismet (Breazeal, 2003) or in virtual agents such as Greta (Bevacqua, Mancini, & Pelachaud, 2004). However, many modern systems (see table below) incorporate both emotion generation and emotional effects, as the two aspects form a loop – either standalone or within cycles of perception and action/expression.

Table 2. Some notable computational models of emotion and cognition, their fundamental theoretical traditions, and effects modeled.

Model	Base Cognitive Theory	Emotion Theory	Effects Modeled
ACRES/WILL (Moffat, Frijda, & Phaf, 1993)	BDI, Planning, Decision Theory, Agents	Appraisal: Frijda	Coping: goal shift, attention shift
ActAffAct (Rank, 2009)	Agents, BDI, Unified Cognition	Appraisal: Frijda, Scherer	Coping: choice of Relational Action Tendency
ACT-R ext. (Cochran, Lee, & Chown, 2006)	ACT	Arousal, valence, clarity: Chown, Damasio. Kleinsmith & Kaplan 1963	Base activation of any memory decays over time if encoded with low arousal, grows with high arousal
ACT-R ext. (Fum & Stocco, 2004)	ACT	Arousal, valence; Damasio 1994, Bechara 2000	Memories associated with risk have higher activation strength for emotion-enabled agents
ACT-R ext. (Belavkin, 2001)	ACT	Arousal, valence; Yerkes-Dodson 1908	Negative valence aids problem solving process, up to a certain level of arousal
(Ahn, 2010)	BDI Motivation, urges, arousal, valence; Loewenstein & Lerner 2003	Reinforcement learning biases (anticipatory reward)	
ALEC (Gadanho & Custodio, 2003)	CLARION	Appraisal: Sloman, Damasio 1994	Decision rules learned based on past experience
EM (Reilly & Bates, 1992)	Oz architecture	Appraisal: OCC; Reilly and Bates 1992	Plan change
EMA (Marsella & Gratch, 2009)	BDI, Agents, Decision Theory, Planning (Newell/Soar)	Appraisal: Smith & Lazarus, Scherer; Simon 1967, Lazarus 1990	Coping: attention shift, plan changes, BDI changes, action tendency changes
Émile (Gratch, 2000)	Strips Planning	Appraisal: OCC; Sloman 1992 et al.	Plan change, plan selection criteria

Table 3. Some other notable computational models of emotion and cognition, their fundamental theoretical traditions, and effects modeled.

Model	Base Cognitive Theory	Emotion Theory	Effects Modeled
EM-ONE (Singh, 2005)	Minsky-Sloman	Appraisal; Minsky 2006, Sloman 2001	Modification of "narratives": plans, desires, or beliefs. Modification of "critic" processes
FAtiMA (Dias, Mascarenhas, & Paiva, 2011)	BDI	Appraisal: OCC; Sloman 1992; Lazarus 1991	Coping: plan and goal changes
FLAME (El-Nasr, Yen, & Ioerger, 2000)	Planning, decision theory, Q-Learning	Appraisal: OCC, Roseman; Bolles and Fanslow 1980, LeDoux 1996	Choice and inhibititon of plans, emotion-based learning and conditioning
(Gmytrasiewicz & Lisetti, 2002)	Decision Theory, Agents	Appraisal: OCC, Frijda, Scherer; Simon 1967	Alotted planning time changes, state utility shifts, state evocation probability shifts
H-CogAff (Sloman, 2001)	BDI, Cognition and Affect (Sloman)	Appraisal: Sloman, OCC; Simon 1967, Sloman 1996	Attention shift (alarms), decision biases, precognitive reactions
MAMID (Hudlicka, 2007)	Belief Net, Decision Theory	Appraisal: Scherer, Smith & Kirby, Sloman; Ortony et al. 2005	Biases mental constructs (data) based on emotional state; Working memory capacity, speed; attention shift, inference speed and biases
(Meyer, 2006)	KARO	Appraisal (OCC); LEA (Logic of Emotional Agents); Meyer 2006	Plan/agenda changes; Fear causes cautious planning
(Malfaz & Salichs, 2006)	BDI, Q-learning	Motivation: Lorentz; Rolls 2003	Reinforcement learning biases (both encoding and recall)
Soar-Emote (Marinier, Laird, & Lewis, 2009)	Soar, PEACTIDM (Newell)	Appraisal: Roseman, Scherer; Gross & John 2003	Attention shift, goal shift, reinforcement learning biases (both encoding and recall)
Tabasco (Petta, 2003)	ACT, BDI	Appraisal (Leventhal & Scherer, Lazarus, Smith et al.)	Plan updates
WASABI (Becker-Asano & Wachsmuth, 2009)	BDI	PAD; Gratch & Marsella 2001, Oatley & Johnson-Laird 1987	Plan utility valuation process biased towards optimism or pessimism, mapping of emotions as beliefs, action biases

Although their designs are widely varied, the systems above collectively illustrate several desirable principles for a comprehensive computational model of emotion and cognition. One is a well-defined base cognitive theory or integrated cognitive process model. For instance, EMA and Soar-Emote both use the Soar cognitive architecture, and there have been several adaptations of the

ACT-R cognitive system that incorporate emotional interaction with the ACT model of memory. Many successful models of emotion and cognition also spring from a strong theoretical emotional background, allowing fine-grained and consistent emotional evaluation parameters.. Note that only a few of the systems above, however, use a unified theory of cognition and emotion for both parts of their model: H-CogAff, EM-ONE, and Meyer's work; of those systems, only EM-ONE was ever implemented. The importance of this observation is in a question of primacy and integration of purpose. Systems built on an integrated cognition, for instance, the Soar and ACT-R based systems, load emotional cognitive operators and memory mechanisms for the modeling of a very specific set of emotional effects. On the other hand, other systems have sophisticated emotional theoretical bases, but are built using BDI or other simple models of cognition. We believe that a system with a broad, integrated cognitive and emotional base model has the most power for modeling and explaining the interactions between human emotion and cognition.

The cognitive emotional effects modeled by systems outlined in Tables 2 and 3 can be categorized as biases or heuristics; similar in scope, but detrimental or useful, respectively, depending on the agent's environment and situation. For instance, attention/focus shift is commonly modeled. Systems that address this include MAMID (Hudlicka, 2007) and H-CogAff (Sloman, 2001). One of the sequential modules of MAMID is devoted to cognitive attention focus, which biases the system toward a subset of incoming data for further processing. H-CogAff (Sloman, 1996), similarly, has an oversight mechanism for sensing pattern-driven "alarms" from all levels of its cognitive processing (reactive, deliberative, and reflective). This mechanism redirects the system to process the stimulus that invoked the alarm.

Effects are often cast as constraints on goal and action choices (i.e., decisions), though there are other types of effects as well. The effects represent a form of coping in EMA (Marsella & Gratch, 2009) and Émile (Gratch, 2000), among other systems. Emotion can affect Émile's planning algorithm so that, for example, the more intensely emotional elements are focused on. In EMA, appraisal and coping are interdependent in a closed loop, and the strategy for building a plan to cope with a particular emotional stimulus is subject to change following the next round of appraisal. Meyer's (2006) system takes a different approach: emotions cause global effects on search control during planning; for instance, a sad agent is more likely to look for alternative plans or goals, whereas a fearful agent will be cautious and perform more checks on its environment during planning and execution.

Decision biases (for planning or otherwise) are also characterized by Becker-Asano's (2009) WASABI and Rank's ActAffAct (2009). In WASABI, the agent's overall emotional state ("core affect" or mood) constrains the set of possible next actions and goals. BehBehBeh and other models of Frijda's theory such as ACRES/WILL (Moffat, Frijda, & Phaf, 1993), use the concept of Relational Action Tendencies (RATs) in a similar manner to constrain decisions; RATs are formed as a direct result of appraisals and narrow the set of next action choices. ALEC (Gadanho & Custodio, 2003) uses a fast emotional system that operates asynchronously along with its cognitive system to model the Somatic Marker Hypothesis during decisions.

Emotional biases on learning are typically memory-based, and may be used to reinforce recall and decision biases. Memories are evaluated as, or become associated with, particular emotional experiences; cognitive effects follow from these evaluations and associations. Recent work by Hyungil Ahn (2010), in addition to modeling decision under emotional influence, also leverages an agent's previous emotional experience for predictive purposes using prospect theory. The result is fast, subjective reinforcement learning, and decision biases result from previous experience. FLAME (El-Nasr, Yen, & Ioerger, 2000) uses a fuzzy logic method for similar purposes, conditioning an agent by mapping emotional states to remembered events.

Emotion as a recall heuristic has been handled in different ways by the systems that have modeled it. ACT-R, with its well-tested model of associative memory, has been a natural starting point for these systems. Fum and Stocco's ACT-R extension (2004), for example, takes advantage of ACT-R's associative memory to reproduce the Iowa Gambling Task's results (though with a skeptical view towards the Somatic Marker Hypothesis). MAMID also models emotional effects on cognitive recall and inference, particularly changes to the speed and capacity of those processes based on emotional appraisal.

The columns in Table 4 are based on the intuitively important interactions between emotion and cognition. Collectively, these pieces combine to form the full loop of emotional and cognitive interaction. Cognitive architectures have found much benefit to separating processes from knowledge, so we follow suit here. The columns for cognitive representation and emotion representation indicate whether a model contains components that are rich enough to allow both cognitive and emotional processes to operate over that knowledge. Examples of limitations of each can be found in the discussion of landmark systems in section 3. The columns "Cognition $\rightarrow$ Emotion" and "Emotion $\rightarrow$ Cognition" represent the full breadth of interaction between cognition and emotion, one for each direction. For an example, a model that addresses how emotion influences inference, learning, decision making, and all other cognitive processes found in the literature would be rated as complete. Discussion of this assessment may be lengthy and we could not find a way to include it, so we defer it to another paper.

A model with complete fidelity consistent with psychological data would receive full rating in all categories. A four star rating represents competence to meet (theoretically as well as experimentally) a wide variety of requirements but not all, three stars competence in meeting maybe a focused group of related requirements, two stars competence in meeting maybe one requirement (e.g., explain one phenomena), one star competence meeting a portion of some requirements, and no stars meaning it is completely unaddressed.

Assuming more refinement may be needed in this assessment, it still should be capable of suggesting that the role of traditional cognition in the generation of emotions is seemingly most lacking. We also could not find any system that had demonstrated, even theoretically, full competency in any area, indicating that considerable research is still needed in this area.

3. Example Models

To understand the state of the art, we review three different computational models in depth that differ greatly in what we may learn from them. These models were selected based on their prominence in the literature and because they demonstrate different aspects of the relationship between emotion and cognition. In each case, we examine the ways in which aspects of cognition contribute to the

Table 4. Rating of each model's competency in key components of interaction between cognition and emotion. A system that addresses all needs for a component of interaction would be rated with ⋆⋆⋆⋆⋆.

Model	Cognitive Representation	Cognition → Emotion	Emotion Representation	Emotion → Cognition
ACRES/WILL (Moffat, Frijda, & Phaf, 1993)	⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆
ActAffAct (Rank, 2009)	⋆⋆	⋆⋆⋆⋆	⋆⋆	⋆⋆⋆
ACT-R extension (Cochran, Lee, & Chown, 2006)	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆	⋆⋆⋆⋆
ACT-R extension (Fum & Stocco, 2004)	⋆⋆⋆⋆	⋆	⋆⋆	⋆⋆⋆
ACT-R extension (Belavkin, 2001)	⋆⋆⋆⋆	⋆⋆	⋆⋆	⋆⋆⋆
(Ahn, 2010)	⋆⋆	⋆	⋆	⋆⋆⋆⋆
ALEC (Gadanho & Custodio, 2003)	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆	⋆⋆⋆
EM (Reilly & Bates, 1992)	⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆
EMA (Marsella & Gratch, 2009)	⋆⋆⋆⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆
Émile (Gratch, 2000)	⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆
EM-ONE (Singh, 2005)	⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
FearNot!/FAtiMA (Dias, Mascarenhas, & Paiva, 2011)	⋆⋆	⋆⋆⋆	⋆⋆	⋆⋆⋆
FLAME (El-Nasr, Yen, & Ioerger, 2000)	⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
(Gmytrasiewicz & Lisetti, 2002)	⋆⋆⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
H-CogAff (Sloman, 2001)	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
MAMID (Hudlicka, 2007)	⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
(Meyer, 2006)	⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆⋆
(Salichs & Malfaz)	⋆⋆	⋆⋆	⋆⋆	⋆⋆⋆
Soar-Emote (Marinier, Laird, & Lewis, 2009)	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆
Tabasco (Petta, 2003)	⋆⋆⋆⋆	⋆⋆⋆⋆	⋆⋆⋆	⋆⋆⋆
WASABI (Becker-Asano & Wachsmuth, 2009)	⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆

generation of emotions and how these emotions are used in cognition. We seek to discuss some of the interesting things we may learn from this approach and identify shortcomings.

3.1 EMA

EMA (Marsella & Gratch, 2009) is a system, implemented on top of the Soar cognitive architecture, as a series of cognitive operators, that is primarily concerned with explaining the dynamics of emotions through a sequence of events.

The primary mechanism used to display this is the use of appraisals informing planning. Appraisal frames are created and associated with plan steps, leading to emotion derivation (or coping directly), then coping. They were the first major system to demonstrate the effective use of an appraisal frame in such a way. The appraisal process is considered a distinct process that is to coordinate with cognitive processes, though most details of how is left unanswered. They also illustrated their perspective on a long standing debate about emotional processes and their timing issues with cognition. Marsella and Gratch argued and showed within EMA that the order and temporal patterns in which appraisal information is provided may just be an artifact of the natural unfolding of the cognitive or physical process which is called. For an example, inference may be slower than recall and appraisals which depend on inference will simply be provided after inference is complete.

The mechanism of appraisal over a plan space is a primary contribution of this work. What we read from it is how classical planning and building of expectations and the violation of confirmation of those expectations can lead to very specific emotion dynamics. However, effects on the two primary cognitive activities of planning and inference (both of which exist in the Soar system) have not been demonstrated in EMA.

Another limitation comes with their use of appraisal frames. Though appraisal frames are by far the most powerful representation of emotions and allow for the most sophisticated effects to be modeled, only a few simple configurations of up to three variables have been demonstrated to have consequence while an appraisal frame in EMA has six variables (argued as minimal). Each variable has a continuous range of values but can also be null, leaving for arbitrarily many configurations. Those who seek to study more emotional effects on cognition as observed in the psychological, cognitive science, and neuroscience literature would need to find on their own how to use these collections of appraisal frames.

Although Marsella and Gratch argue that the discrete emotion labels derived from the appraisal frames are merely for convenience, they encode rules that map appraisal configurations to specific coping behaviors as well as to discrete emotion labels. It is arguable that this is equivalent to translating an appraisal frame into a discrete emotion and then to a set of coping responses. This is similar to many other approaches (e.g., WASABI, as discussed below) used which seem to lose granularity of context when eliciting emotional response.

3.2 Soar-Emote (PEACTIDM)

Soar-Emote (Marinier, Laird, & Lewis, 2009) is another system built on top of the Soar cognitive architecture, in which emotions result from PEACTIDM, an explicit model of cognitive control (Newell, 1990) and Scherer's appraisal theory. This effort included a proposal to integrate an emotional component at the Soar architectural level in the form of an appraisal detector. Emotions are also represented in appraisal frames, but unlike EMA there is really only one appraisal frame of significance in the system per cognitive cycle.

The two seemingly most interesting contributions of this work are how PEACTIDM may interact with emotions and how different appraisals influence the notion of intensity. The PEACTIDM work demonstrates how appraisals may be integrated in the cognitive cycle. The proposals for using appraisal information to calculate intensity answers some questions about how to calculate both arousal and valence in various models. This also suggests a mechanism for appraisal information to be implicit in a simple valence/arousal value which has been demonstrated to affect many forms of cognition in other work.

Soar-Emote proposes appraisal activity as being distinct from traditional cognitive processes; a shortcoming as per EMA. Though Marinier agrees that translating an appraisal frame into a discrete emotion or a PAD value does not make sense, as every possible appraisal frame should elicit a different response, there are no proposed details for how this can be done, leaving many open questions. Its designers have, however, shown great examples for how this model would play in a maze sample problem as well some suggestions for how emotion may affect learning (Marinier & Laird, 2008).

3.3 WASABI

WASABI (Becker-Asano & Wachsmuth, 2009) may be one of the most general models of emotion that has been built to believably simulate affective agents. This was one of the few models built from the ground up with emotions as first class citizens within the design.

Several interesting lessons may be gleaned from WASABI, particularly from the modeling of primary emotions and secondary emotions. Many theories try to explain the connection between emotions that seem to be instantaneous and emotions that are more complex and seem to arise from reasoning. In contrast to EMA which proposes this as a dynamic development of emotions in time through a reappraisal process, WASABI suggests grounding of secondary emotions in primary in-born emotions.

WASABI offers an example of how many other systems (e.g., FAtiMA) approach integration of emotion and cognition. Though appraisals are made, this information is mostly translated into either discrete emotions or, in the case of WASABI, a PAD space value to be used in affecting cognitive processes. As mentioned in the discussion of EMA, this approach is a lossy compression of information, losing all the context and specifics of the emotion.

4. Open Issues

As computational models of emotion and cognition and their integration are still a relatively new topic of research, the open issues are broad and numerous. Here we intend to draw attention to what we believe are the important issues for computational emotions and cognition at large.

The most discernible issue is that of criteria and methods for model evaluation. This has been a difficult problem to crack in the area of cognitive architecture and, since emotion likely has an intimate relationship with nearly all components of cognitive architecture, this problem among others has carried over. At the forefront of this issue are Gratch and Marsella, who have proposed several methods of evaluation: compare the model in its ability to perform like a human in a standard clinical assessment for emotions and coping (Gratch & Marsella, 2005), encode a corpus of emotional situations within a model and compare results (Gratch, Marsella, & Mao, 2006), and, most recently, a component framework for comparing models (Marsella, Gratch, & Petta, 2010). At an abstract level, we would ideally evaluate a model by comparing its ability to explain all observed human emotional phenomena. However, the breadth of phenomena associated with emotion research is so vast that different models tend to focus on very different phenomena, typically demonstrating value but leaving little common ground for comparison. Encoding a corpus of emotional situations would provide a powerful way of testing if the details of model are consistent with that of human processes and data. However, this approach currently suffers from the issues of psychological theory, in that the emotional situations, their encoding, and associated details are products of inference and interpretation. The component model allows comparisons between systems and highlights theoretical similarities and differences, but, provides little aid in showing validity.

To address this, we suggest the development of a library of standard scenarios to test models. This idea is being explored by those tackling the evaluation of cognitive architectures, with some proposals made by Adams et al. (2012). A similar approach may be used for emotional phenomena. This differs slightly from the previously mentioned approach by Gratch et al. (2006) in that a standardized environment would be provided as the starting point, rather than an encoded emotional situation, which may not provide direct comparability with other models.

The second issue, which has also carried over from cognitive architecture research, is that of the domain's complexity and breadth, and the inability to effectively reduce it. The literature that informs emotion research crosses the domains of computer science, psychology, cognitive science, behavioral economics, sociology, and others. Minsky (2007) also characterizes the human brain as an extremely sophisticated system of systems and argues that, to model human intelligence with fidelity, we cannot shy away from its complexity . Important aspects of emotions come in its broad integration with various components of cognition. This means that many computational emotion researchers must deal with many topics for which they cannot possibly hold enough expertise to develop complete models. To manage the complexity, some researchers have followed the lead of other artificial intelligence sub-disciplines and studied emotions separately from cognitive architecture. We believe this approach is flawed since it is becoming increasingly clear that emotions are fundamentally intertwined with all forms of cognition. Research in computational emotion cannot use the same approaches from most other AI fields and must be studied in full context of cognition.

The previous issue naturally leads into a third, which can be further broken down into smaller ones of note, that concerns implementing a model in the full context of a cognitive architecture. This is a notoriously difficult task in a research setting, since a cognitive architecture is typically a large and complex system that requires many contributors over a long period of time. When integrating emotions into an architecture, one must leverage the rich history of AI research on individual topics like learning, planning, and reasoning, but it is nearly impossible for any researcher to be an expert in emotions and all of these topics. There is also a question of how to coordinate among all the processes, as there are several theories and incomplete data about timing. This leaves many details to assume, which can lead to flaws that may be hard to understand, particularly as validation often focuses on emergent behavior–a product of all processes, data, and design decisions. Some previous models have tried to jump this hurdle by building accounts of emotions on top of an existing cognitive architecture, either as a by-product of the overall system or as a small modification to one mechanism. However, emotions appear to have developed early in the brain's evolutionary journey; they have a clear presence in the reptilian brain, and thus predate all forms of higher cognition (Panksepp, 2001). As a result, these approaches have given us some insight but have yet to reveal the fundamental nature of emotions and how cognition operates as their consequent.

The second issue and third issues, taken together, are unavoidably large and complex and require cross-disciplinary collaboration. In response, we encourage the creation of a collaborative community based around an open flexible architecture. This endeavor is ambitious but it would help alleviate the greatest hurdle for such systems–the need for broad expertise, long development time, and a large code base. Although all architectures require certain theoretical commitments (e.g., knowledge representation), modern engineering techniques can help abstract these away, making it possible for individuals and small research groups to replace them. If we can make a cognitive architecture framework that is flexible and reconfigurable, then various labs can experiment with very different architectures but leverage work done by others, letting them focus on specific problems within their expertise. This would address the remaining issues regarding complexity, implementation, and realization of actual architectures, as well as integration with the rich history of AI. Using modern open source and crowd sourcing approaches, we can also build a library of modules focused on specific cognitive processes (e.g., learning, metacognition). A similar effort known as OpenCog (Goertzel et al., 2010) has been underway for some time now, but it is primarily concerned with applications rather than scientific experimentation and understanding.

5. Challenges to Near-Term Research

Aside from the high level open issues, there are several important challenges to research in emotions and cognition that warrant more immediate attention from researchers. These include moving towards uniformity in emotional representations and mechanisms, understanding existing use of emotions in traditional artificial intelligence, exploring innovative uses of emotions, and emotion engineering. These issues are raised in hopes of steering the current direction of research and we discuss their implications and their potential to drive forward work on computational emotion.

We have pointed out the intimate interplay between emotion and cognition found in psychological and neurological literature, but there is a clear deficit in models that explicitly study this relationship. Several systems are partly competent at modeling some interplay, but they suffer from either being narrowly focused on one mechanism of interaction (e.g., Ahn, 2010; Fum & Stocco, 2004) or involve a model that engineered the relationship as an afterthought to another goal (e.g., believable emotional behavior in WASABI). As emotions are intertwined with all forms of cognition, explicit study and modeling in the general context of a cognitive architecture should be fundamental to further understanding and using them. To achieve this, future research should be geared towards uniformity of emotional representations and mechanisms. Analogously, Rosenbloom (2009) has argued that representational uniformity is the key to integration of a broad and diverse set of capabilities required for general intelligence.

A common question that is raised by researchers in a subfield of artificial intelligence, when presented with ideas for how emotions can aid in their processes, is: "How is that different from *item X* that we already have?" Examples of such topics include heuristics used in plan space search, positive/negative feedback for reinforcement learning, and expertise for action selection. From a narrow perspective, it does not appear that emotions provide any additional benefit to artificial intelligence; however, from a cognitive systems perspective, if all these things are linked to emotions, then they can provide uniform data across processes that explain how this information is generated. To date, there has been no work on what artificial intelligence and can learn from models of emotion, but we believe that researchers interested in both areas would benefit.

Since research on emotion is relatively new, there is opportunity for it to unlock answers to difficult problems in artificial intelligence. Exploration of emotions and its various uses has potential to suggest novel responses to open problems. For example, creativity in the arts (e.g., music, painting, story writing) often involves emotional expression, understanding, and communication. Understanding emotions' role in the creative arts may teach us something that can be generalized to other forms of creativity.

Finally, it is important for researchers in cognitive systems to be able to produce systems with specific behaviors. With emotions serving as general cognitive data with broad influence, there is the challenge of engineering emotional knowledge to these desired behaviors. For example, some emergent qualities such as personality appear to arise from affect (Arnold, 1960; Asensio et al., 2008). Traditional knowledge engineering principals, methods, and tools involving representation, human driven knowledge authoring, understanding and maintenance should be developed to support computational research on emotion.

6. Conclusion

In this paper, we reviewed the landscape of research on computational models of emotion and cognition. We analyzed their levels of emotional-cognitive integration to help understand how each system compares and contrasts with others. We also identified several key properties of the models, and we suggested that researcher strove for total competency in each area in order to fully model emotions. We described and analyzed three models–EMA, Soar-Emote, and WASABI–in depth to provide a broad sample of the state of the art, and we identified the most significant contributions and limitations of each system.

In addition, we identified the significant open issues that warrant most attention from researchers. These included standardizing criteria for evaluation of models, the complexity and breadth of the domain, and the need to implement working systems that address integration with the rich history of AI research. We then presented and discussed possible responses to this challenge. The community should provide standard scenarios with known behavioral results to contextualize models, and it should leverage collaborative approaches to software engineering to let experts in each field contribute and experiment.

For the near term, we posed four challenges that could steer research in computational emotion, including a focus on uniform representations and mechanisms, understanding the role of emotions in subfields of artificial intelligence, exploring innovative uses of emotions, and engineering knowledge bases.

References

Adams, S. S., Arel, I., Bach, J., Coop, R., Goertzel, B., Hall, J. S., Samsonovich, A., Scheutz, M., Schlesinger, M., Shapiro, S. C., & Sowa, J. (2012). Mapping the landscape of human-level artificial general intelligence. *AI Magazine*, *33*, 25–42.

Ahn, H.-I. (2010). *Modeling and analysis of affective influences on human experience, prediction, decision making, and behavior*. Doctoral dissertation, Department of Architecture, Massachusetts Institute of Technology, Cambridge, MA.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

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

Asensio, J., Jimenez, M., Fernandez, S., & Borrajo, D. (2008). A social and emotional model for obtaining believable emergent behaviors. *Artificial Intelligence: Methodology, Systems, and Applications; Lecture Notes in Computer Science*, *5253*, 395–399.

Bechara, A., Damasio, H., & Damasio, A. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.

Becker-Asano, C., & Wachsmuth, I. (2009). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems*, *20*, 32–49.

Belavkin, R. (2001). The role of emotion in problem solving. *Proceedings of the AISB'01 Symposium on Emotion, Cognition and Affective Computing* (pp. 49–57). York, UK.

Bevacqua, E., Mancini, M., & Pelachaud, C. (2004). Speaking with emotions. *Proceedings of the AISB'04 Symposium on Motion, Emotion and Cognition* (pp. 197–214). University of Leeds, UK.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition*, 59–82. Combridge University Press.

Bower, G. (1983). Affect and cognition. *Philosophical Transactions of the Royal Society of London*, *B*, 387–402.

Breazeal, C. (2003). Toward sociable robots. *Robots and Autonomous Systems*, *42*, 167–175.

Cochran, R., Lee, F., & Chown, E. (2006). Modeling emotion: Arousal's impact on memory. *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1133–1138).

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.

Dias, J., Mascarenhas, S., & Paiva, A. (2011). FAtiMA Modular: Towards an agent architecture with a generic appraisal framework. *Proceedings of the International Workshop on Standards for Emotion Modeling*.

El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). FLAME - Fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent Systems*, *3*, 219–257.

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

Fum, D., & Stocco, A. (2004). Memory, emotion, and rationality: An ACT-R interpretation for gambling task results. *Proceedings of the Sixth International Conference on Cognitive Modeling* (pp. 106–111).

Gadanho, S., & Custodio, L. (2003). Asynchronous learning by emotions and cognition. *Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior* (pp. 224–225).

Gmytrasiewicz, P. J., & Lisetti, C. L. (2002). Emotions and personality in agent design and modeling. In S. Parsons, P. Gmytrasiewicz, & M. Wooldridge (Eds.), *Game theory and decision theory in agent-based systems*, 81–95.

Goertzel, B., Garis, H. D., Pennachin, C., & Geisweiller, N. (2010). OpenCogBot: Achieving generally intelligent virtual agent control and humanoid robotics via cognitive synergy. *Proceedings of the Twelfth Annual International Conference on Artificial Intelligence* (pp. 1–12).

Gratch, J. (1996). Why you should buy an emotional planner. *Proceedings of the Agents'99 Workshop on Emotion-based Agent Architectures* (pp. 99–107).

Gratch, J. (2000). Émile: Marshalling passions in training and education. *Proceedings of the Fourth International Conference on Autonomous Agents* (pp. 325–332).

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

Gratch, J., & Marsella, S. (2005). Evaluating a computational model of emotion. *Autonomous Agents and Multi-Agent Systems*, *11*, 23–43.

Gratch, J., Marsella, S., & Mao, W. (2006). Towards a validated model of "emotional intelligence". *Proceedings of the 21st National Conference on Artificial Intelligence* (pp. 1613–1616). Menlo Park, CA: AAAI Press.

Gratch, J., Marsella, S., & Petta, P. (2009). Modeling the cognitive antecedents and consequences of emotion. *Cognitive Systems Research*, *10*, 1–5.

Hudlicka, E. (2007). Reasons for emotions: Modeling emotions in integrated cognitive systems. In W. Gray (Ed.), *Integrated models of cognitive systems*, 1–37. New York: Oxford University Press.

Laird, J. E. (2008). Extending the Soar cognitive architecture. *Proceedings of the 2008 Conference on Artificial General Intelligence* (pp. 224–235). Amsterdam.

Langley, P., Laird, J., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*, 141–160.

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

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*, 1019–1024.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lehrer, J. (2009). *How we decide*. Boston: Houghton Mifflin Harcourt.

Leventhal, S., & Scherer, K. R. (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and Emotion*, *1*, 3–28.

Malfaz, M., & Salichs, M. (2006). Using emotions for behaviour-selection learning. *Proceedings of the 17th European Conference on Artificial Intelligence* (pp. 697–698). Amsterdam: IOS Press.

Marinier, R., & Laird, J. (2008). Emotion-driven reinforcement learning. *Cognitive Science*, 115–120.

Marinier, R., Laird, J., & Lewis, R. (2009). A computational unification of cognitive behavior and emotion. *Cognitive Systems Research*, *10*, 48–69.

Marsella, S. C., & Gratch, J. (2009). EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, *10*, 70–90.

Marsella, S. C., Gratch, J., & Petta, P. (2010). Computational models of emotion. In K. R. Scherer, T. Banziger, & E. Roesch (Eds.), *A blueprint for affective computing: A sourcebook and manual*, 21–45. New York: Oxford University Press.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.

Meyer, J.-J. C. (2006). Reasoning about emotional agents. *International Journal of Intelligent Systems*, *21*, 601–619.

Minsky, M. (2007). *The Emotion Machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. New York: Simon and Schuster.

Moffat, D., Frijda, N., & Phaf, R. (1993). Analysis of a model of emotions. In A. Sloman, D. Hogg, G. Humphreys, & A. Ramsay (Eds.), *Prospects for artificial intelligence*, 219–228. Amsterdam: IOS Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge, MA: Cambridge University Press.

Panksepp, J. (2001). The neuro-evolutionary cusp between emotions and cognitions: Implications for understanding consciousness and the emergence of a unified mind science. *Evolution and Cognition*, *7*, 141–163.

Petta, P. (2003). The role of emotions in a tractable architecture for situated cognizers. In R. Trappl, P. Petta, & S. Payr (Eds.), *Emotions in humans and artifacts*, 251–288. Cambridge, MA: MIT Press.

Rank, S. (2009). *Behaviour coordination for models of affective behaviour*. Doctoral dissertation, Vienna University of Technology, Vienna.

Reilly, W. S., & Bates, J. (1992). *Building emotional agents* (Technical Report). School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Rosenbloom, P. S. (2009). Towards uniform implementation of architectural diversity. *Proceedings of the AAAI Fall Symposium on Multi-Representational Architectures for Human-Level Intelligence* (pp. 32–33). Arlington, VA: AAAI Press.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Scherer, K. R. (1999). Appraisal theory. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion*, 637–663. Chichester, UK: John Wiley & Sons.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*, 92–120. New York: Oxford University Press.

Schorr, A. (2001). Appraisal: The evolution of an idea. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research*, 20–33. New York: Oxford University Press.

Singh, P. (2005). *EM-ONE: An architecture for reflective commonsense thinking*. Doctoral dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

Sloman, A. (2001). Varieties of affect and the CogAff architecture schema. *Proceedings of the AISB'01 Symposium on Emotion, Cognition, and Affective Computing* (pp. 39–48). York, UK.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.