Affective Computing
Ph.D. Comprehensive Exam

Mohammad Shayganfar - mshayganfar@wpi.edu
May, 26 2015

1 Introduction to Affective Computing

According to Picard [64], the term affective computing encapsulates a new approach in Artificial Intelligence, to build computers that show human affection. Studies show that the decision making of humans is not always logical [31], and in fact, not only is pure logic not enough to model human intelligence, but it also shows failures when applied in artificial intelligence systems [19]. Emotions impact fundamental parts of cognition including perception, memory, attention and reasoning [15]. This impact is caused by the information emotions carry about the environment and event values.

If we want robots and virtual agents to be more believable and efficient partners for humans, we must consider the personal and social functionalities and characteristics of emotions; this will enable our robots to coexist with humans, who are emotional beings. To have a better understanding of applications of affective computing, we can categorize the whole existing literature of computational emotion modeling and their applications into four major categories of: a) detecting and recognizing human emotions, b) interpreting and understanding human emotions, c) generating artificial emotions and applying the underlying processes to exploit emotion functions, and d) expressing human-perceivable emotions during interaction.

There are some major emotion theories including *appraisal, dimensional* and *discrete (basic)* theories, some of which have corresponding computational models, e.g., EMA [55] and WASABI [8, 9]. These models have been used in different domains including AI and robotics. Modeling and applying these theories can help robots and virtual agents to achieve communicative, evaluative, interpretive, and regulatory aspects of emotions in some or all of the four application domains we mentioned above.

This document provides description of major computational emotion theories, their comparison, and their applications in AI and robotics. It includes the existing influential computational emotion theories as well as the underlying psychological theories; it majorly focuses on appraisal and dimensional theories, although it briefly mentions other approaches, e.g. discrete (basic) emotions.

2 Computational Theories of Emotion/Affect

There are different types of computational theories of emotion. These theories differ in the type of relationships between their components and whether a particular component plays a crucial role in an individual emotion. For instance, the basic component of an emotion can be the behavioral tendencies, the cognitive elements, or the somatic processes. Emotion theories can also differ based on their representationial distinction.

2.1 Appraisal Theory

Appraisal theories of emotion were first formulated by Arnold [5] and Lazarus [45] and then were actively developed in the early 80s by Ellsworth and Scherer and their students [71] [74] [75] [79] [81]. The emotional experience is the experience of a particular situation [25]. Appraisal theory describes the cognitive process by which an individual evaluates the situation in the environment with respect to the individual's well-being and triggers emotions to control internal changes and external actions.

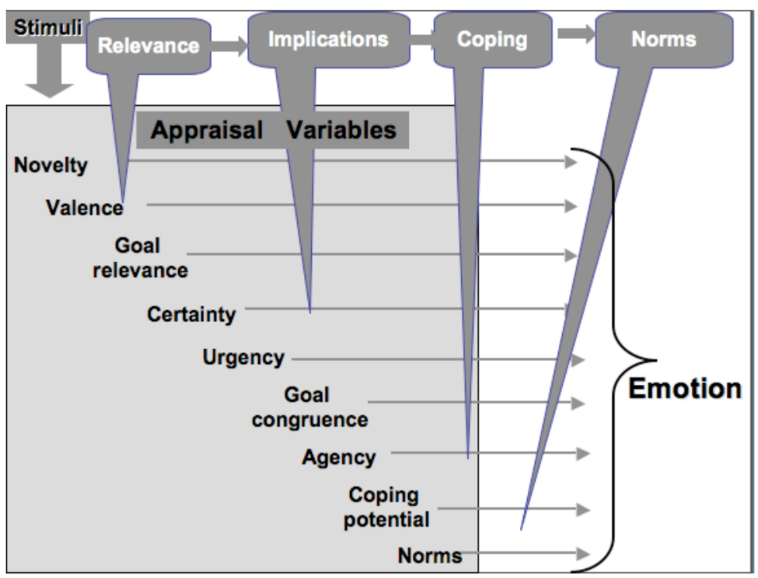

Figure 1: Schematic view of the componential theory of emotion [38].

2.1.1 Componential Approach

This approach emphasizes the distinct components of emotions, and is often called the *componential* approach [47]. The "components" referred to in this approach are the components of the cognitive appraisal process. These are referred to as *appraisal variables*, and include *novelty, valence, goal relevance, goal congruence*, and *coping abilities* (further on, in this section, some of the appraisal variables used in computational models are introduced) [75, 81]. A stimulus, whether real or imagined, is analyzed in terms of its meaning and consequences for the agent, to determine the affective reaction. The analysis involves assigning specific values to the appraisal variables. Once the appraisal variable values are determined by the organisms evaluative processes, the resulting vector is mapped onto a particular emotion, within the n-dimensional space defined by the n appraisal variables. The semantic primitives for representing emotions within this model are thus these individual appraisal variables. Figure 1 shows the relationship of the individual appraisal dimensions to the broader categories of evaluations taking place during appraisal (Relevance, Implications, etc.).

2.1.2 Component Process Model

The Component Process Model (CPM) is Scherer's influential and major theory of emotions [76, 81]. This theory focuses on the dynamic unfolding of emotions. The CPM suggests that an event and its consequences are appraised with a set of criteria on multiple levels of processing (the appraisal component). The result of the appraisal will generally have a motivational effect, often changing or modifying the motivational state before the occurrence of the event. Based on the appraisal results and the motivational changes, some effects will occur in the autonomic and somatic nervous system. The CPM considers emotions as the synchronisation of many different cognitive and physiological components. Emotions are identified with the overall process whereby low level cognitive appraisals, in particular the processing of relevance, trigger bodily reactions, behaviours and subjective feelings. The model suggests that there are four major appraisal objectives required to adaptively react to a salient event [78]:

a) **Relevance:** How relevant is this event for the agent? Does it directly affect the agent or its social reference group?

b) **Implications:** What are the implications or consequences of this event and how do they affect agent's well-being and its immediate or long-term goals?

c) **Coping Potential:** How well can the agent cope with or adjust to these consequences?

d) **Normative Significance:** What is the significance of this event for the agent's self-concept and for social norms and values?

To attain these objectives, the agent evaluates the event and its consequences on a number of criteria or *Stimulus Evaluation Checks* (SECs), with the results reflecting the agents subjective assessment of consequences and implications on a background of personal needs, goals, and values [81]. Figure 2 shows the postulated sequence, the cognitive and motivational inputs and the effects on response systems. Also, the bidirectional effects between appraisal and other cognitive functions are illustrated by the arrows in the upper part of Figure 2.

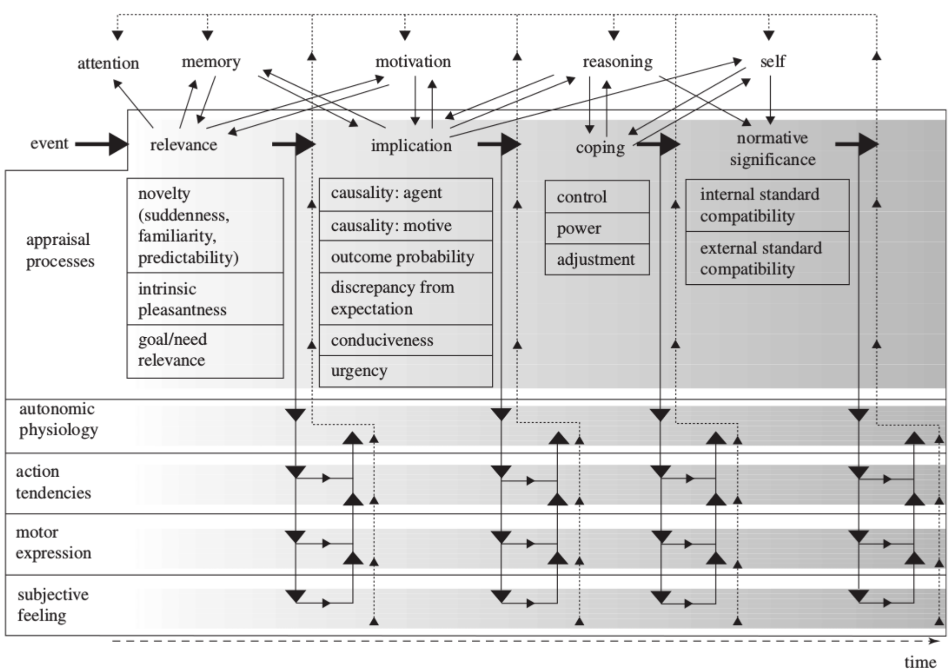

Figure 2: Comprehensive illustration of the CPM of emotion [78, 81].

2.1.3 Appraisal Process

According to this theory, appraisals are separable antecedents of emotion, that is, the individual first evaluates the environment and then feels an appropriate emotion [81]. The appraisal procedure begins with the evaluation of the environment according to the internalized goals and is based on systematic assessment of several elements [76]. The outcome of this process triggers the appropriate emotions. In many versions of appraisal theory, appraisals also trigger cognitive responses often called *coping strategies*. In fact, the coping mechanism manages the individual's action with respect to the individual's emotional state and the existing internal and/or external demands [24]. The large majority of computational models of emotions are based on this theory. An individual can also use knowledge about the emotional reactions of others to make inferences about them. According to the appraisal patterns, different emotions can be experienced and expressed. Since expression of emotions reflects one's intentions through the appraisal process, the *reverse appraisal* mechanism helps one to infer others' mental states based on their expressions. [18, 33].

Appraisal process is typically viewed as the cause of emotion and the cognitive and behavioral changes associated with emotion. For instance, a particular pattern of the appraisal variables (i.e., individual judgements) will elicit a certain emotion or emotional expressions. These appraisal variables include [55]:

- **Relevance:** A relevant event has non-zero utility for an agent. This relevancy can either be based on a negative influence of an event on the agent or a positive one.
- **Perspective:** The point of view in which an event will be judged, e.g. self or other.
- **Desirability:** A desirable event advances a state of the utility for an agent whose perspective is being taken, or if it is an undesirable event, inhibits that.
- **Likelihood:** A measure of likelihood of the outcome.
- **Expectedness:** The extent to which the truth value of a state could have been predicted from causal interpretation.
- **Causal Attribution:** The agent who deserves the credit/blame.

- **Controllability:** Whether the outcome can be altered by the agent whose perspective is taken (this variable is related to the coping process).

- **Changeability:** Whether the outcome can be altered by some other causal agent (this variable is related to coping process).

2.1.4 Coping Process

Another key process involved in appraisal is the coping process. This process determines whether and how the agent shoud respond with respect to the outcome of appraising the events. There are several coping strategies that computational models like EMA [29] use as control signals. These control signals enable or suppress the cognitive processes that operate on the causal interpretation of the appraisal patterns. The coping process controls the congruency of the actions according to these patterns. As it is shown below, in [29] coping strategies are organized into two categories: *problem-focused* and *emotion-focused*. Problem-focused coping strategies can be applied when the agent must do something with respect to the problem, whereas Emotion-focused coping works by changing one's interpretation of circumstances. The following is a short list of a broad range of coping strategies [29]:

Problem-focused coping

- **Active coping:** Taking active steps to remove or circumvent the stressor,
- **Planning:** Coming up w/ action strategies,
- **Seeking social support for instrumental reasons:** Seeking advice, assistance, or information.

Emotion-focused coping

- **Seeking social support for instrumental reasons:** Getting sympathy, moral support or understanding,
- **Acceptance:** Accepting the stressor and learning to live with it,
- **Restraint coping:** Waiting till the appropriate opportunity (holding back).

2.1.5 OCC, a Structural Appraisal Theory of Emotion

OCC (Ortony, Clore and Collins) model, similar to Lazarus' [44] and Scherer's [75] cognitive views, considers emotions to arise from affective or valenced reactions subsequent to the appraisal of a stimulus as being beneficial or harmful to ones concern [59]. The model categorizes emotions based on their underlying appraisal patterns. These patterns are fundamental criteria a person employs for evaluating a situation. They involve the person's focus of attention, her concern, and her appraisal preceding an affective reaction. Figure 3 shows main building blocks of OCC model.

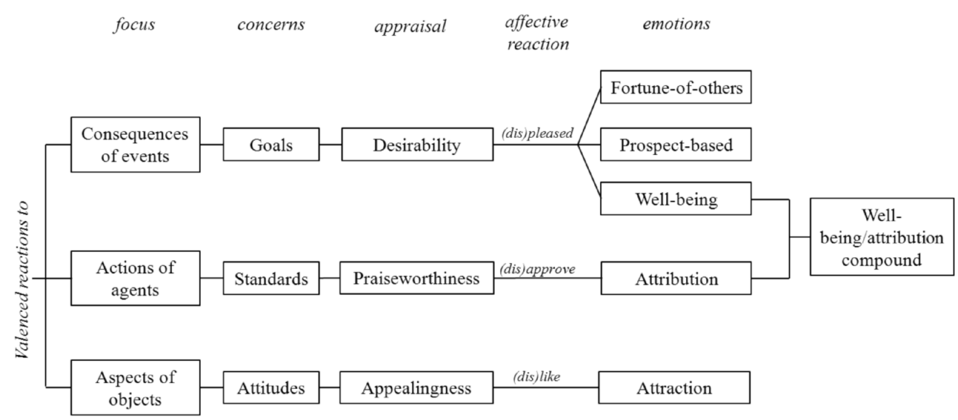

Figure 3: A simple visualization of OCC model [59].

As shown in Figure 3, a person could alternatively have three types of focuses. These types of focuses are consequence of events, actions of agents, and aspects of objects. The person evaluates the significance of causes behind these three types of focuses based on her personal concern. As a result, an affective reaction will be elicited resulting in an emotion. Various combinations of the elements depicted in Figure 3 create specific patterns demonstrating six main groups of emotions in which all emotion types in a group share the same cognitive pattern. Emotion groups are *fortune-of-others, prospect-based, well-being, attribution, well-being/attribution- compound*, and *attraction*. The OCC model introduces 22 emotion types. These emotions are introduced each as a representative of a family of similar emotions with various intensities (since relying on a list of discrete emotions that is understood by everyone equally is impossible due to people's language barriers and various interpretations of the actual words). For instance, happiness can be referred to by other emotion terms such as joy, cheerfulness, gladness, delighted while they all share the same eliciting conditions. Thus the emotion types used in the model (e.g., relief, love, pride, and shame) are meant to represent an emotional experience rather than a lexical taxonomy.

For instance, as shown in Figure 3, the appraisal criterion for consequences of events is their *desirability* (see Section 2.1) for achieving one's goals. This generates the affective reaction of being *pleased* in positive cases, or *displeased* in negative ones. Figure 4 shows the resulting emotion groups in OCC model such as *fortune-of-others* (e.g., gloating, pity), *prospect- based* (e.g., satisfaction, relief), and *well-being* (e.g., joy, distress) [59]. The appraisal of the praiseworthiness of the actions of an agent against one's personal standards, as well as the appealing aspects of objects happens in the same way as shown in Figure 3.

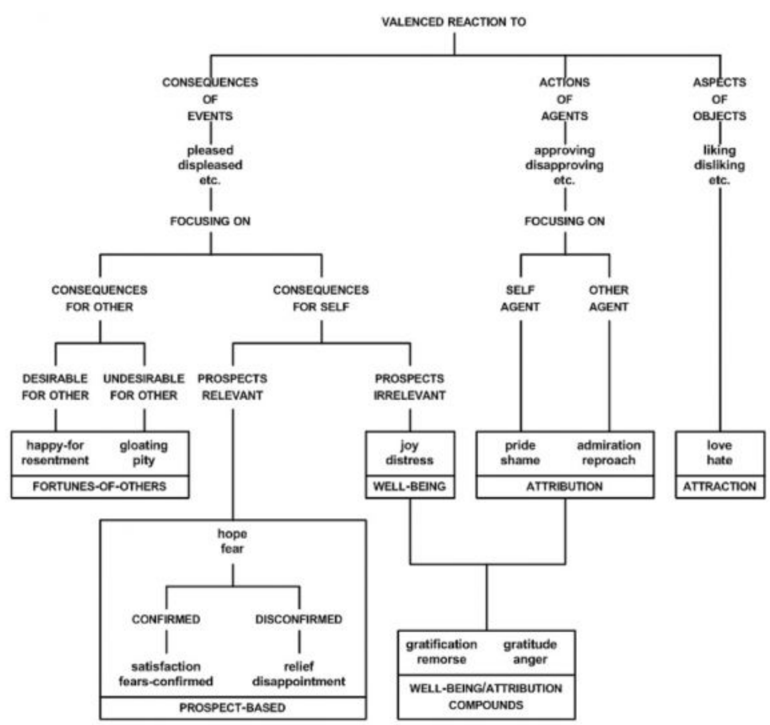

Figure 4: OCC taxonomy of emotion triggers and emotions [59].

Finally, the OCC model introduces some global variables of an emotion's intensity to distinguish all types of emotions that a person could experience when encountering events, agents or objects. These variables are as follows:

1. Sense of reality (representing the degree to which the event, agent or object in focus appear real to the person),
2. Proximity variable (representing the psychological proximity of an event, agent or object),
3. Unexpectedness (representing how surprising an event is for one, either positive or negative),
4. Arousal (representing how arousing an event, agent or object is).

2.2 Constructivist (Dimensional) Emotion Theories

The components and dimensions of emotions were the subject of much speculation since the 19th century. Dimensional models of emotion attempt to conceptualize human emotions by defining where they lie in two or three dimensions. Dimensional theories of emotion argue that emotion should be conceptualized, as points in a continuous (typically two or three) dimensional space rather than looking at them as discrete entities (see Section 2.3) [14] [58] [73] [95].

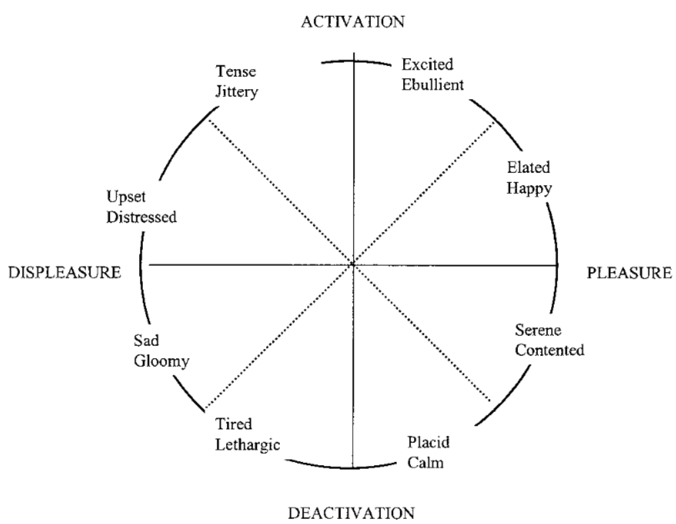

Figure 5: Russell's suggested affective states based on core affect [73].

Two dimensions that are commonly proposed to describe emotions are valence and physiological arousal [5] [44] [72]. Models based on dimensional theories contrast theories of basic emotion, which propose that different emotions arise from separate neural systems [66]. Many dimensional theories argue that discrete emotion categories (e.g., sadness, fear and anger) have no "reality" in that there are no specific brain regions or functions that correspond to specific emotions [7]. Dimensional theories do not emphasize the term emotion.

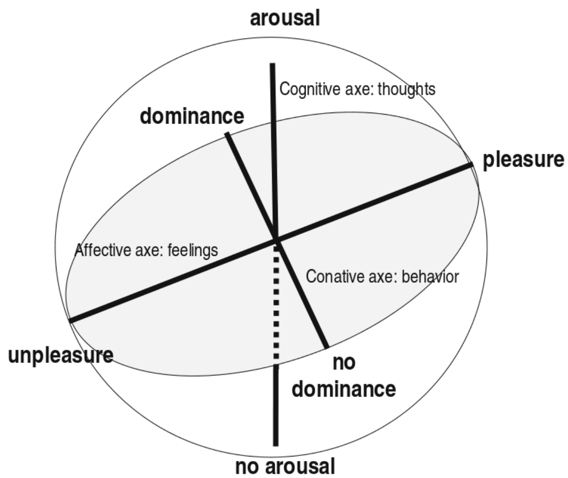

Figure 6: Three dimensional model of pleasure, arousal and dominance as tripartite view of experience [6].

One of the most prominent two-dimensional models is Russell's circumplex model [72]. Russell suggested that affective states are all related to each other systematically through what is called core affect [72, 73] (see Figure 5) and emotions are best described as a change in core affect which, in turn, is describable as a point in a space between two bipolar dimensions. One dimension is *valence* or how good or bad objects and events are for a being ranging from pleasant to unpleasant. The other dimension is *arousal*, ranging from calm to excited. Russell put a number of affective states around a circular space between those two dimensions (see Figure 5) which is also known as *circumplex*, representing the variety of core affects [72, 73]. Since sometimes two-dimensional space cannot easily differentiate among emotions that share the same values of arousal and valence, e.g., anger and fear (both characterized by high arousal and negative valence), some of the dimensional models incorporate valence and arousal as well as *intensity*, or *dominance* or *stance* dimensions. Many computational dimensional models build on the three dimensional PAD model of Mehrabian and Russell [58] where these dimensions correspond to pleasure (a measure of valence), arousal (indicating the level of affective activation) and dominance (a measure of power or control). Figure 6 shows these three dimensions.

Figure 7: Basic emotions and corresponding expressions.

2.3 Basic (Discrete) Emotion Theories

Basic emotion theories are inspired by Tomkins' [90] rediscovery of Darwin's work [17, 35] which later were developed by Ekman [20] and Izard [40]. These theories emphasize a small set of discrete and fundamental emotions. The underlying assumption of this approach is that these emotions are mediated by associated neural circuitry, with a hardwired component [20]. Different emotions are then characterized by stable patterns of triggers, behavioral expression, and associated distinct subjective experiences. The emotions addressed by these theories are typically called the *basic* emotions. Emotions including happiness, sadness, fear, anger, surprise, and disgust are often considered to comprise the most prototypical basic emotions [20]. The theory of basic emotions holds that there is a set of emotions shared by all humans that evolved to deal with ancestral life challenges [20]. For instance, disgust evolved to deal with the challenge of avoiding noxious stimuli, and fear evolved to deal with the challenge of avoiding dangers. Because of the emphasis on discrete categories of states, this approach is also termed the *categorical* approach [63]. Much of the supporting evidence offered for the theory comes from experiments that show how certain facial expressions are universally associated with specific basic emotions, regardless of the observer's cultural background. This universality has a production side and a recognition side. On the production side, a particular emotional state is said to elicit a facial expression comprised of a fixed set of facial muscles. On the recognition side, observers are able to infer the emotional state of the person who expresses an emotion, due to the direct correspondence between emotional states and the facial expressions they cause. Computational models inspired by the basic emotions or discrete approach often focus on low-level perceptual-motor tasks and encode a two-process view of emotion that argues for a fast, automatic, undifferentiated emotional response and a slower, more differentiated response that relies on higher level reasoning processes (e.g., [4]).

2.4 Other Approaches

There are other approaches that different researchers take based on their emphasis on the applicability of emotions in their systems.

2.4.1 Rational Approaches

Rational approaches start from the question of what adaptive functions emotions serve and then attempt to incorporate these functions into a model of intelligence. Emotion, within this approach, is simply another set of processes and constraints that have adaptive value. Models of this sort are most naturally directed towards the goal of improving theories of machine intelligence [3] [82] [85].

2.4.2 Communicative Approaches

Communicative theories of emotion argue that emotion processes function as a communicative system. They can function first, as a mechanism for informing other individuals of ones mental state (thereby facilitating social coordination), and second, as a mechanism for requesting/demanding changes in the behavior of others. Communicative theories emphasize the social-communicative function of expressions [28]. Computational models inspired by communicative theories focus on machinery that decides when an emotional expression can have a desirable effect on a human counterpart.

3 Similarities and Differences

Different theoretical perspectives should not be viewed as competing for a single truth. They should be seen as distinct perspectives, each arising from a particular research area (e.g., biological vs. social psychology), focusing on different sets of affective phenomena, considering distinct levels of resolution and fundamental components (e.g., emotions vs. appraisal variables as the distinct primitives). These different perspectives also provide different degrees of support for the distinct processes of emotion, e.g., the componential theories provide extensive details about cognitive appraisals [38]. Therefore, I am going to provide a pairwise comparison between these fundamental theories. Note that a distinct pairwise comparison will not be provided for appraisal vs. discrete (basic) emotion theories as important points are adequately covered in the comparisons presented below.

3.1 Dimensional Vs. Discrete (Basic) Emotion Theories

The fundamental assumption of the basic emotion theory is that a specific type of event triggers a specific affect program corresponding to one of the basic emotions and producing characteristic expression patterns and physiological response configurations [79]. Dimensional theory's main criticism of basic emotions theory is based on the observation that affective phenomena appear to be both qualitatively and quantitatively diverse.

Russell in [73] argues the labels such as "fear", "anger", "happiness" do not capture this diversity. For instance, one might say: a) a person being chased by an assailant brandishing a knife, b) a person who retreats from an insect moving across the floor, and c) a person who is concerned they will never find a fulfilling career are all in a state of fear. On the basic emotions account, an emotional episode involves fixed patterns of neurophysiological and facial expression changes in response to an eliciting stimulus that are distinct between emotions, but are the same within the same emotional category [20]. If this were the case, one would expect that the three individuals described above would respond to their eliciting stimuli in the same way, yet the similarity of behavioral responses between these three cases seem unlikely. Dimensional theorists, in contrast, would argue that the individuals in the above three cases are applying the concept of fear to experience, despite the fact that each individual has a unique core affect. While basic emotion theorists would hold that since all three individuals are experiencing fear, they would perform the same behavioral responses to the stimuli, dimensional theorists would argue this is not the case, as each individual bears a core affective state that is distinguished from the other two. For instance, the individual's arousal in response to an armed assailant should be higher than the individual in response to an insect, as the former case poses a threat to their life. As a result, the individual in the first case would likely make every effort to escape from the assailant, including trying to negotiate and plead with the assailant, while the individual in the second case would be relatively less dedicated to escaping the insect.

In sum, dimensional theory is compatible with the differences in the behavioral responses to eliciting stimuli, while basic emotions theory only allows for a single fixed behavior of responses to a given emotion. Furthermore, dimensional theories can represent instances of basic emotions (see Figure 8), for example, fear elicited by a snake (green rectangle), in terms of variation along affective dimensions, i.e., arousal and valence.

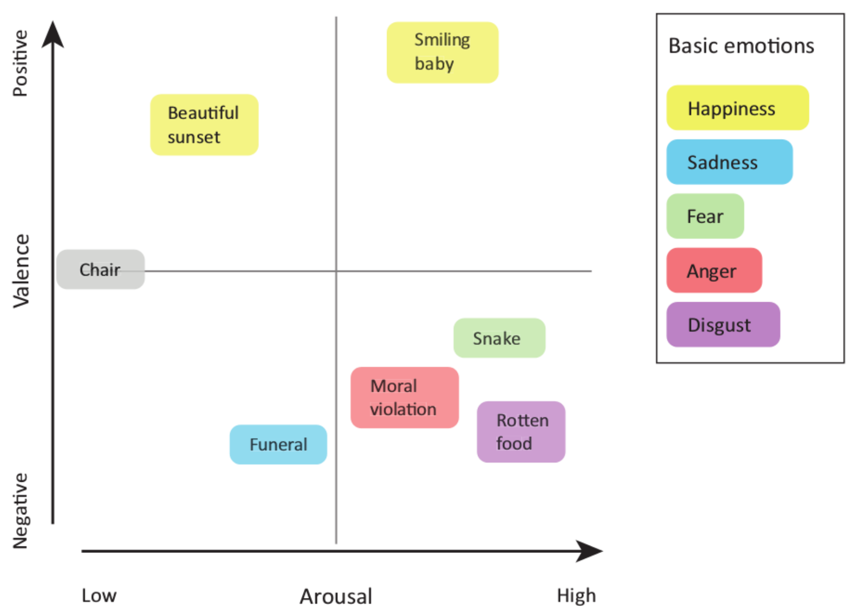

Figure 8: Representing basic emotions within a dimensional framework [32].

Also, basic emotion theory fails to account for affect that lacks object-directedness [73]. In basic emotions approach, an emotion is supposed to have an intentional object it is directed towards (e.g., being angry at someone, or being sad for someone). The dimensional theory argues that emotion may not necessarily be aimed at a particular object. For instance, an individual can experience a certain type of emotion (e.g., anger) without knowing of anything in particular that has offended her. Dimensional models of emotion are therefore capable of accounting for a wider range of affective phenomena than basic emotions theory.

Another difference between dimensional and basic emotion theories is that the basic emotion categorization of emotions captures facets of the experience of an emotion not conveyed by the dimensional description, such as elicitation of a facial expression of the emotion. In fact, this attribute of the basic emotions theory is one of the major differences with all other emotion theories. As it is argued in basic emotion theory, basic emotions are hard-wired to their corresponding facial expressions. Ekman who elaborated the concept of basic emotions, developed the *Facial Action Coding System* (FACS) which encodes movements of individual facial muscles and it is a common standard to systematically categorize the physical expression of emotions [21].

3.2 Appraisal Vs. Dimensional Emotions Theories

Dimensional theories might struggle to adequately distinguish emotions because of the existence of limited dimensions.

To compare the appraisal and dimensional theories of emotion, we can argue that there is a relationship between the dimensions in the constructivist or dimensional theory of emotion and appraisal dimensions. For instance, the pleasure dimension roughly maps onto appraisal dimensions that characterize the valence of an appraisal-eliciting event (e.g., intrinsic pleasantness –desirability–, or goal congruence), dominance roughly maps onto the appraisal dimension of coping potential, and arousal can be considered as a measure of intensity. However, they also have quite different meanings. Appraisal (as I mentioned earlier) is a relational construct characterizing the relationship between some specific object/event in the environment and the individual's mental constructs including beliefs, motives and intentions and several appraisals may be simultaneously active; whereas emotions in dimensional emotion theory are non-relational constructs, each summarizing a unique overall state of the individual.

Furthermore, dimensional emotion theories emphasize different components of emotion than appraisal theories and link these components quite differently. In contrast to appraisal theories, dimensional emotion theories do not address affects antecedents in detail. However, dimensional theorists question the tight causal linkage between appraisal and emotion that is central to appraisal accounts. As mentioned earlier, dimensional theorists believe that the emotion is not necessarily about some object (as in "I am angry <u>at him</u>"). In such theories, many factors may contribute to a change in emotion including intentional judgments (e.g., appraisal). However, in dimensional emotion theories the link between any preceding intentional meaning and emotion is broken and most of the time can not be recovered correctly. For example, Russell argues for the following sequence of emotional components: some external event occurs (e.g., a bear walks out of the forest), it is perceived in terms of its affective quality; this perception results in a crucial change in core affect; this change is attributed to some "object" (e.g., the bear); and only then is the object cognitively appraised in terms of its goal relevance, causal antecedents and future prospects [54].

We can also compare the dimensional emotion theories to OCC model as a cognitive appraisal model. The major similarity between these two models is that they both consider emotions to descend from valenced reactions to the stimuli. Furthermore, they acknowledge the role of arousal in determining emotional reactions. As we mentioned in Section 2.2 Russell considered arousal as one of the two key dimensions of emotions which could be used to partially discriminate emotional states [72]. In a different manner, the OCC model recognizes arousal as a necessary condition for eliciting emotions, and regards the arousal as a major determinant of the elicited emotion's intensity which distinguishes among various emotions of a particular type (e.g., fearful and scared). In [77] Scherer speculates that the arousal dimension in dimensional models gives little information about the underlying appraisal of the elicited emotion and he proposes to replace it with coping potential which is an appraisal dimension referring to the individual's perceived control in a given situation.

Furthermore, models based on dimensional emotions theory pursue the idea of eliciting an emotion according to the joint features in circumplex space (2D or 3D – see Section 2.2) while OCC or other models of appraisal theory are based on patterns of antecedents of emotions. This is the fundamental difference between OCC, or appraisal theories in general, and the circumplex approach of Russell [72] or Mehrabian's PAD model [6, 58]. Also, models based on appraisal theory of emotion employ causation, attribution and eliciting conditions in order to distinguish emotions while the eliciting conditions are not directly accessible from dimensional approach. A dimensional model might fall short in establishing why certain emotions are elicited. However, when the objective is to identify the generated emotions and their level of pleasantness and intensity, a circumplex model presents a perfect opportunity [2].

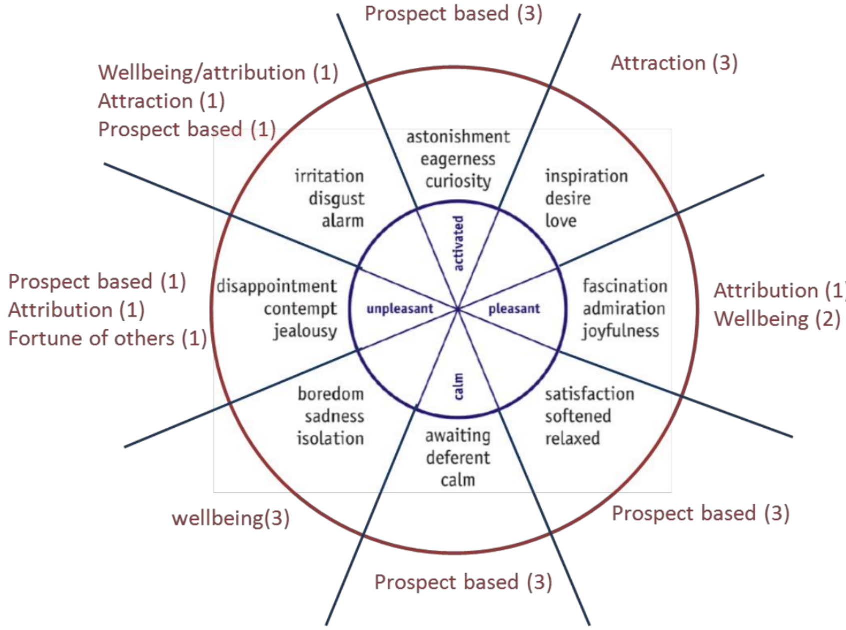

Figure 9: A rough projection of emotion groups of OCC on the circumplex of affect [2].

Finally, here, I discuss how a model based on dimensional emotions theory (i.e., Russell's 2D circumplex) relates to a cognitive model based on appraisal theory (i.e., OCC model). Figure 9 shows the relationship between Russell's circumplex and OCC model in terms of categorization of the actual emotions. The number of emotions in a section of Russell's circumplex that fall into an emotion group of OCC are shown in parentheses (see Figure 9). For instance, all three emotions in the top section (highly excited, neutrally valenced emotions) fall into prospect based emotion group, hence number (3) is indicated. Or, as another example, emotions in the left section (neutral arousal value, negative valenced emotions) make a one to one relationship between disappointment and prospect based emotion group, contempt and attribution emotion group, and jealousy and fortune of others emotion group, hence number (1) is indicated in front of each.

4 Applications in Autonomous Agents and Robots

There are many research areas, including robotics and autonomous agents, that employ the structure and/or functions of emotions in their work with a variety of motivations behind modeling emotions [96]. Some of these works are inspired by specific psychological theories (we provide several examples in this section), some are freely using the concept of emotion without using the theoretical background in social sciences, and some are using a combination of concepts from the psychological theories. For instance, in PECS [92] which is designed for modeling human behaviors, the agent's architecture is not based on a certain kind of social or psychological emotion theory. In fact, it is intentionally designed and described in a way which enables the integration of a variety of theories. The PECS' design enables an integrative modeling of physical, emotional, cognitive and social influences within a component-oriented agent architecture. Also, in [56] the computational architecture which is designed to provide information about the possible overall behavior of a work team is not based any specific theory. As we mentioned earlier, some researchers apply combinations of emotion theories in their work [42]. For instance, in [13] Cañamero shows how an agent can use emotions for activity selection while taking into account both dimensional and discrete approaches in an action selection mechanism. Through out this section, we provide different examples of works using major emotion theories in robots and autonomous agents.

We can also see the application of emotion theories in designing companion robots, robots capable of expressing emotions and social behaviors, as well as robots which can convey certain types of emotion products, e.g., empathy [12] [46] [62] [83]. Robots also use emotions theories for automatic affect recognition using different modalities [34] [97]. Moreover, in some works, researchers have explored the users affective state as a mechanism to adapt the robot's behaviors during the interaction [11] [50].

Applications of Appraisal Theory – The emphasis of models derived from appraisal theories of emotion is on making appraisal the central process. Computational appraisal models often exploit elaborate mechanisms for deriving appraisal variables such as decision-theoretic plans [29] [55], reactive plans [68] [70] [87], Markov-decision processes [22] [84], or detailed cognitive models [52]. However, emotion itself is sometimes treated less elaborately, and simply as a label to which behavior can be attached [23]. Appraisal is usually modeled as the cause of emotion being derived via simple rules on a set of appraisal variables.

Computational appraisal models have been applied to a variety of uses including contributions to psychology, robotics, AI, and HCI. For instance, Marsella and Gratch have used EMA [55] to generate specific predictions about how human subjects will appraise and cope with emotional situations and argue that empirical tests of these predictions have implications for psychological appraisal theory [30] [53]. There are several examples in artificial intelligence and robotics of applying appraisal theory [1] [41] [55]. In robotics, appraisal theory has been used to establish and maintain a better interaction between a robot and a human. For instance in [41] researchers provide their computational model of emotion generation based on appraisal theory to have a positive human-robot interaction experience. In [74] authors describe a system approach to appraisal processes based on Scherer's work on appraisal and the Component Process Model [75]. They show how the temporal unfolding of emotions can be experimentally tested. They also lay out a general domain-independent computational model of appraisal and coping. In [94] researchers consider their robot's (INDIGO) emotion, speech and facial expressions as a key point to establish an effective communication between the robot and a human during their interaction. They apply concepts of appraisal theory in INDIGO's emotion modeling. MAGGIE, a sociable robot, also applies the appraisal theory of emotions to consider fear in its decision making system [27]. Velasquez developed Cathexis which is a distributed computational model for generation of emotions and their influence in the behavior of the autonomous agents [93]. The emotion model in this work is based on Roseman's work on appraisal theory. Marinier and Laird in [39] focus on the functional benefits of emotion in a cognitive system. In this work, they integrate their emotion theory (which is based on the appraisal theory) with Soar cognitive architecture, and use emotional feedback to drive reinforcement learning. In [37] Hudlicka provides a model of a generic mechanism mediating the affective influences on cognition based on cognitive appraisal. This model is implemented within a domain-independent cognitive-affective architecture (MAMAID).

In the virtual agents community, empathy is a research topic that has received much attention in the last decade [10] [57] [60] [67] [88]. In [65] researchers developed an agent with capability of affective decision-making based on appraisal theory to establish an affective relationship with its users. Then, they compared the performance of their agent with a human (based on a WoZ study) in a speed-dating experiment. In HCI, the appraisal theory has been primarily used for the creation of interactive characters that exhibit emotions in order to make characters more believable [69], more realistic [51] [91], more capable of understanding human motivational states [16] or more able to induce desirable social effects in human users [61].

Applications of Dimensional Theory – The emphasis of models influenced by dimensional theories is on processes associated with core affect which is usually represented as a continuous time-varying process, and it can be determined at a given time by a point in a 2D or 3D-space as a response to the eliciting events. Generally, there are detailed mechanisms in computational dimensional models which determine how this point changes over time, e.g., decay to some resting state, and incorporating the impact of dispositional tendencies such as personality or temperament [26] [54]. The models based on dimensional theories have also been used in robotics. For instance, researchers in [48] apply PAD's three-dimensional space to rate the pleasure, arousal and dominance of their Multimodal Emotional Intelligence robot (MEI) in each interaction with human subjects. Their goal is to introduce the first steps in MEI which can understand and express emotions in voice, gesture and gait. In [98] researchers want to understand the effect of different interface features for a service robot. They use valence and arousal dimensions in their questionnaires to assess the perceived anthropomorphism of their own service robot by their subjects. In [43] researchers introduce the implementation of a dynamic personality for a robot based on a dimensional emotion model. They use WASABI's architecture [8, 9] as their emotional model. In [49] the author describes an affective knowledge representation scheme to be used in the design of a socially intelligent artificial agent. Lisetti uses the valence-arousal two dimensional model of emotion in this work. This model has been applied in an emotion-based architecture of Lisetti's autonomous robots as well as a multimodal affective user interface agent. ROMAN, an expressive robotic head, uses a behavior-based emotional control architecture. The approach to the emotional component of the architecture is based on the dimensional emotion theory [36].

Comparison of Applications of Emotion Theories – Researchers often use computational dimensional models for behavior generation of animated characters. The reason might be because it is easier for emotion translation to a limited number of dimensions that can be readily mapped to continuous features of behavior such as the spatial extent of a facial expression. For example, PAD models describe all behavior in terms of only three dimensions of pleasure, arousal and dominance, whereas researchers using appraisal models should either associate each behavior with a large number of appraisal variables [80] [86], or try to map appraisal variables into a limited and small number of discrete expressions [23]. For a similar reason, dimensional models also frequently used as a good representational framework for systems that attempt to recognize human emotional behavior and there is some evidence that they may better discriminate user affective states than approaches that rely on discrete labels [7].

There is also a relationship between dimensional and appraisal theories. Some of the computational models of emotion that incorporate dimensional theories have viewed appraisal as the mechanism that initiates changes to core affect. For instance, ALMA [26] includes OCC inspired appraisal rules [59], and WASABI [8] includes appraisal processes inspired by Scherer's sequential-checking theory into a PAD-based emotion model. Moreover, some computational models explore how core affect in dimensional models can influence cognitive processes. For example, HOTCO 2 [89] allows explanations to be biased by dimensional affect [54].

5 Conclusion

In this response, we started by looking at the description of affective computing and the importance of the concept of emotion in general. Then, we provided our four categories of computational models of emotions which we can consider for the applications of affective computing.

There are major theories of emotions explaining the concept of emotion. We discussed these major theories in detail separately, providing their psychological background and underlying concepts. Following the explanation of these theories, we were able to discuss the similarities and differences between these major theories. Finally, we provided applications of these theories in robotics and AI.

We believe to develop or work based on computational models of emotions, it is good to follow well-established (in comparison with others) theoretical foundations. These theories can be a guideline for our computational models, and they can explain more details of the structure or the processes involved in affective situations. However, we do not necessarily think that the computational models must exactly follow only one theory and its descriptions. Meaning, different aspects of models can represent different theories. For instance, appraisal theory is a good representation of the interpretive aspect of emotions and basic emotion theories provide detailed systematic methods for expressive application. More importantly, we believe the interpersonal functions of emotions should be our first concern and try to relate them to the structure of our domain, i.e., collaboration. In conclusion, we can see the importance of interpretive, communicative and regulatory aspects of emotion functions in this proposed work.

References

[1] Carole Adam and Emiliano Lorini. A BDI emotional reasoning engine for an artificial companion. In *Workshop on Agents and multi-agent Systems for AAL and e-HEALTH (PAAMS)*, volume 430, pages 66–78. Springer, 2014.

[2] Naseem Ahmadpour. Occ model: application and comparison to the dimensional model of emotion. In *In Proceedings of International Conference on KANSEI Engineering and Emotion Reseach*, pages 607–617, 2014.

[3] John R. Anderson and Christian Lebiere. The newell test for a theory of cognition. *Behavioral and Brain Sciences*, 26(5):587–640, 2003.

[4] Armony, Servan-Schreiber, Cohen, and Ledoux. Computational modeling of emotion: Explorations through the anatomy and physiology of fear conditioning. *Trends in Cognitive Sciences*, 1:28–34, 1997.

[5] Magda B. Arnold. *Emotion and personality.* Cassell Co., 1960.

[6] Iris Bakker, Theo van der Voordt, Peter Vink, and Jan de Boon. Pleasure, arousal, dominance: Mehrabian and russell revisited. *Current Psychology*, 33(3):405–421, 2014.

[7] Lisa Feldman Barrett. Are emotions natural kinds? *Perspectives on Psychological Science*, 1(1):28–58, 2006.

[8] Christian Becker-asano. *WASABI: Affect Simulation for Agents with Believable Interactivity.* PhD thesis, University of Bielefeld, 2008.

[9] Christian Becker-Asano. Wasabi for affect simulation in human-computer interaction: Architecture description and example applications. In *Proceedings of ERM4HCI workshop in conjunction with ICMI 2013*. Springer, 2013.

[10] Scott Brave and Clifford Nass. Emotion in human-computer interaction. In Julie A. Jacko and Andrew Sears, editors, *The Human-computer Interaction Handbook*, pages 81–96. L. Erlbaum Associates Inc., 2003.

[11] Cynthia Breazeal. *Designing Sociable Robots.* MIT Press, 2002.

[12] Cynthia Breazeal. Role of expressive behaviour for robots that learn from people. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1535):3527–38, 2009.

[13] Lola D. Canamero. Designing emotions for activity selection in autonomous agents. In Robert Trappl, Paolo Petta, and Sabine Payr, editors, *Emotions in Humans and Artifacts*, pages 115–148. MIT Press, 2003.

[14] Charles S. Carver. Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5(4):345–356, 2001.

[15] Gerald L. Clore and Jeffrey R. Huntsinger. How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11(9):393–399, 2007.

[16] Cristina Conati and Heather Maclare. Evaluating a probabilistic model of student affect. In *7th International Conference on Intelligent Tutoring Systems*, 2004.

[17] Charles Darwin. *The expression of the emotions in man and animals.* New York; D. Appleton and Co., 1916.

[18] Celso M. de Melo, Jonathan Gratch, Peter Carnevale, and Stephen J. Read. Reverse appraisal: The importance of appraisals for the effect of emotion displays on people's decision-making in social dilemma. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci)*, 2012.

[19] Hubert L. Dreyfus. *What Computers Still Can'T Do: A Critique of Artificial Reason.* MIT Press, 1992.

[20] Paul Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200, 1992.

[21] Paul Ekman and Wallace V. Friesen. Measuring facial movement. *Environmental psychology and nonverbal behavior*, 1(1):56–75, 1976.

[22] Magy Seif El-Nasr, John Yen, and Thomas R. Ioerger. Flame: Fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-Agent Systems*, 3(3):219–257, 2000.

[23] Clark Davidson Elliott. *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System.* PhD thesis, Northwestern University Institute for the Learning Sciences, 1992.

[24] Susan Folkman and Judith T. Moskowitz. Coping: Pitfalls and promise. *Annual Review of Psychology*, 55:745–774, 2004.

[25] Nico H. Frijda. *The Emotions.* Cambridge University Press, 1986.

[26] Patrick Gebhard. A layered model of affect. In *5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 29–36, 2005.

[27] Alvaro Castro Gonzalez, Maria Malfaz, and Miguel Angel Salichs. An autonomous social robot in fear. *IEEE Transactions Autonomous Mental Development*, 5(2):135–151, 2013.

[28] Jonathan Gratch. True emotion vs. social intentions in nonverbal communication: Towards a synthesis for embodied conversational agents. In *ZiF Workshop*, volume 4930, pages 181–197. Springer, January 2008.

[29] Jonathan Gratch and Stacy C.Marsella. A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.

[30] Jonathan Gratch, Stacy Marsella, Ning Wang, and Brooke Stankovic. Assessing the validity of appraisal-based models of emotion. In *International Conference on Affective Computing and Intelligent Interaction.* IEEE, 2009.

[31] S. Grossberg and W. E. Gutowski. Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review*, 94(3):300–318, 1987.

[32] Stephan Hamann. Mapping discrete and dimensional emotions onto the brain: controversies and consensus. *Trends in Cognitive Sciences*, 16(9):458–466, 2012.

[33] Shlomo Hareli and Ursula Hess. What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition & Emotion*, 24(1):128–140, 2009.

[34] Frank Hegel, Torsten Spexard, Britta Wrede, Gernot Horstmann, and Thurid Vogt. Playing a different imitation game: Interaction with an empathic android robot. In *Proceedings of 2006 IEEE-RAS International Conference on Humanoid Robots (Humanoids06)*, 2006.

[35] Ursula Hess and Pascal Thibault. Darwin and emotion expression. *American Psychologist*, 64(2):120–128, 2009.

[36] Jochen Hirth, Norbert Schmitz, and Karsten Berns. Emotional architecture for the humanoid robot head roman. In *IEEE International Conference on Robotics and Automation*. IEEE, 2007.

[37] Eva Hudlicka. Reasons for emotinos: Modeling emotinos in integrated cognitive systems. In Wayne D. Gary, editor, *Integrated Models of Cognitive Systems*, volume 59, pages 1–37. New York: Oxford University Press, 2007.

[38] Eva Hudlicka. Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions*, 2(1):85–102, 2011.

[39] Robert P/ Marinier III and John E. Laird. Emotion-driven reinforcement learning. In *CogSci 2008*, 2008.

[40] Carol Ellis Izard. *Human Emotions*. NY: PLenum, 1977.

[41] Hyoung-Rock Kim and Dong-Soo Kwon. Computational model of emotion generation for human-robot interaction based on the cognitive appraisal theory. *Journal of Intelligent and Robotic Systems*, 60(2):263–283, 2010.

[42] Kiril Kiryazov, Robert Lowe, Christian Becker-Asano, and Tom Ziemke. Modelling embodied appraisal in humanoids : Grounding pad space for augmented autonomy. In *Proceedings of the Workshop on Standards in Emotion Modeling*, 2011.

[43] Marius Klug and Andreas Zell. Emotion-based human-robot-interaction. In *IEEE 9th International Conference on Computational Cybernetics (ICCC)*, 2013.

[44] Richard S. Lazarus, J. R. Averill, and E. M. Opton. Toward a cognitive theory of emotions. In *Feelings and Emotions*, pages 207–232. New York: Academic Press, New York, 1970.

[45] Richard Stanley Lazarus. *Emotion and Adaptation*. OXFORD University Press, 1991.

[46] Iolanda Leite, André Pereira, Samuel Mascarenhas, Carlos Martinho, Rui Prada, and Ana Paiva. The influence of empathy in human-robot relations. *International Journal of Human-Computer Studies*, 71(3):250–260, 2013.

[47] Howard Leventhal and Klaus Scherer. The relationship of emotion to cognition. *Cognition and Emotion*, 1:3–28, 1987.

[48] Angelica Lim and Hiroshi G. Okuno. The mei robot: Towards using motherese to develop multimodal emotional intelligence. *IEEE Transactions Autonomous Mental Development*, 6(2):126–138, 2014.

[49] Christine L. Lisetti. Personality, affect and emotion taxonomy for socially intelligent agents. In *Proceedings of the Fifteenth International Florida Artificial Intelligence Research Society Conference*. AAAI Press, 2002.

[50] Changchun Liu and Nilanjan Sarkar. Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE TRANSACTIONS ON ROBOTICS*, 24(4):883–896, 2008.

[51] Wenji Mao and Jonathan Gratch. Evaluating a computational model of social causality and responsibility. In *Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, 2006.

[52] Robert P. MarinierIII, John E. Laird, and Richard L. Lewis. A computational unification of cognitive behavior and emotion. *Cognitive System Research*, 10(1):48–69, March 2009.

[53] Stacy Marsella, Jonathan Gratch, Ning Wang, and Brooke Stankovic. Assessing the validity of a computational model of emotional coping. In *International Conference on Affective Computing and Intelligent Interaction*. IEEE, 2009.

[54] Stacy Marsella, Jonathan Grath, and Paolo Petta. Computational models of emotion. In Etienne Roesch Klaus R. Scherer, Tanja Banziger, editor, *A Blueprint for Affective Computing: A Sourcebook and Manual*, pages 21–41. Oxford University Press, 2010.

[55] Stacy C. Marsella and Jonathan Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, March 2009.

[56] Juan Martínez-Miranda, Arantza Aldea, and René Bañares-Alcántara. Simulation of work teams using a multi-agent system. In *The Second International Joint Conference on Autonomous Agents & Multiagent Systems AAMAS, July 14-18, Melbourne, Victoria, Australia*, pages 1064–1065, 2003.

[57] Scott W. McQuiggan and James C. Lester. Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, 65(4):348–360, 2007.

[58] Albert Mehrabian and James A. Russell. *An approach to environmental psychology*. The MIT Press, 1974.

[59] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[60] Ana Paiva, Joao Dias, Daniel Sobral, Ruth Aylett, Polly Sobreperez, Sarah Woods, Carsten Zoll, and Lynne Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 194–201, 2004.

[61] Ana Paiva, Joao Diasa, Daniel Sobrala, Ruth Aylettb, Sarah Woodsc, Lynne Halld, and Carsten Zoll. Learning by fealing: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, 19:235–266, 2005.

[62] Ana Paiva, Iolanda Leite, and Tiago Ribeiro. Emotion modeling for sociable robots. In Jonathan Gratch Arvid Kappas Rafael A. Calvo, Sidney D'Mello, editor, *Handbook of Affective Computing*, pages 296–308. Oxford University Press, 2014.

[63] Jaak Panskepp. *Affective Neuroscience: The Foundations of Human and Animal Emotions*. NY:Oxford University Press, 1998.

[64] Rosalind W. Picard. *Affective Computing*. The MIT Press, 2000.

[65] Matthijs Pontier and Johan F. Hoorn. How women think robots perceive them - as if robots were men. In *International Conference on Agents and Artificial Intelligence (ICAART-2)*, pages 496–504, 2013.

[66] JONATHAN POSNER, JAMES A. RUSSELL, and BRADLEY S. PETERSON. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, null(3):715–734, 2005.

[67] Helmut Prendinger and Mitsuru Ishizuka. The empathic companion: a character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3-4):267–285, 2005.

[68] Stefan Rank and Paolo Petta. Appraisal for a character-based story-world. In *5th International Working Conference on Intelligent Virtual Agents*, Kos, Greece, 2005. Springer.

[69] W. Scott Neal Reilly. *Believable Social and Emotional Agents*. PhD thesis, Carnegie Mellon University, 1996.

[70] W. Scott Neal Reilly. Modeling what happens between emotional antecedents and emotional consequents. In *Eighteenth European Meeting on Cybernetics and Systems Research*, 2006.

[71] Ira J. Roseman and Craig A. Smith. Appraisal theory:overview, assumptions, varieties, controversies. In Klaus R. Scherer, Angela Schorr, and Tom Johnstone, editors, *Appraisal process in emotion*, pages 3–34. NY: Oxford University Press, 2001.

[72] James A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[73] James A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1):145–172, 2003.

[74] David Sander, Didier Grandjean, and Klaus R. Scherer. A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18(4):317–352, 2005.

[75] Klaus R. Scherer. On the nature and function of emotion: A component process approach. In Klaus R. Scherer and Paul Ekman, editors, *Approaches To Emotion*, pages 293–317. Lawrence Erlbaum, Hillsdale, NJ, 1984.

[76] Klaus R. Scherer. On the sequential nature of appraisal processes: Indirect evidence from a recognition task. *Cognition & Emotion*, 13(6):763–793, 1999.

[77] Klaus R. Scherer. What are emotions? and how can they be measured? *Social Science Information*, 44:695–729, 2005.

[78] Klaus R. Scherer. The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23(7):1307–1351, 2009.

[79] Klaus R. Scherer. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3459–3474, 2009.

[80] Klaus R. Scherer and Heiner Eligring. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7(1):113–130, 2007.

[81] Klaus R. Scherer, Angela Schorr, and Tom Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research.* Oxford University Press, 2001.

[82] Matthias Scheutz and Aaron Sloman. Affect and agent control: Experiments with simple affective states. In *Intelligent Agent Technology*, pages 200–209. World Scientific, 2001.

[83] Mohammad Shayganfar, Charles Rich, and Candace L. Sidner. A design methodology for expressing emotion on robot faces. In *IROS*, pages 4577–4583. IEEE, 2012.

[84] Mei Si, Stacy C. Marsella, and David V. Pynadath. Modeling appraisal in theory of mind reasoning. *Autonomous Agents and Multi-Agent Systems*, 20(1):14–31, 2010.

[85] Herbert A. Simon. Motivational and emotional controls of cognition. *Psychology Review*, 74(1):29–39, 1967.

[86] Craig A. Smith and Heather S. Scott. A componential approach to the meaning of facial expressions. In James A. Russell and Jos Miguel Fernndez-Dols, editors, *The psychology of facial expression. Studies in emotion and social interaction*, pages 229–254. NY: Cambridge University Press, 1997.

[87] Alexander Staller and Paolo Petta. Introducing emotions into the computational study of social norms: A first evaluation. *Journal of Artificial Societies and Social Simulation*, 4:615–625, 2001.

[88] Milind Tambe. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.

[89] Paul Thagard. Why wasn't o. j. convicted: emotional coherence in legal inference. *Cognition and Emotion*, 17(3):361–383, 2003.

[90] Silvan S. Tomkins. *Affect, Imagery, Consciousness.* NY: Springer, 1962.

[91] David Traum, Jeff Rickel, onathan Gratch, and Stacy Marsella. Negotiation over tasks in hybrid human-agent teams for simulation-based

training. In *International Conference on Autonomous Agents and Multiagent Systems*, 2003.

[92] Christoph Urban. Pecs: A reference model for human-like agents. In *Deformable Avatars*. Netherlands: Kluwer Academic Publishers, 2001.

[93] Juan D. Velàsquez. Modeling emotions and other motivations in synthetic agents. In *Proceedings of the 14th Nnational Conference on Artificial Intelligence AAAI-97*, pages 10–15, 1997.

[94] Dimitrios Vogiatzis, Constantine Spyropoulos, Vangelis Karkaletsis, Zerrin Kasap, Colin Matheson, and Olivier Deroo. An affective robot guide to museums. In *Proceedings of the 4th International Workshop on Human-Computer Conversation*, 2008.

[95] David Watson and Auke Tellegen. Toward a consensual structure of mood. *Psychological Bulletin*, 98(2):219–235, 1985.

[96] Thomas Wehrle. Motivations behind modeling emotional agents: Whose emotion does your robot have?, 1998.

[97] Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. A survey of affect recognition methods: Audio, visual and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, 2009.

[98] Tao Zhang, David B. Kaber, Biwen Zhu, Manida Swangnetr, Prithima Mosaly, and Lashanda Hodge. Service robot feature design effects on user perceptions and emotional responses. *Intelligent Service Robotics*, 3(2):73–88, 2010.