Affective Computing Models for Character Animation
Ricardo L. Parreira Duarte, Prof. Abdennour El Rhalibi, Christopher Carter, Prof. Madjid Merabti
R.L.Duarte@2010.ljmu.ac.uk; {A.Elrhalibi; C.J.Carter; M.Merabti}@ljmu.ac.uk
School of Computing and Mathematical Sciences
Liverpool John Moores University,
Byrom Street, L3 3AF, Liverpool, UK

Abstract- “No aspect of our mental life is more important to the quality and meaning of our existence than emotions” [1]. Emotions have been one of the most intriguing subjects in the study of human psyche, which gave birth to numerous theories since the ancient Greek philosophers all throughout history. Emotions play an intrinsic role within human social relations and behaviours. However they are also intangible, and until today there is no proven model that defines them fully, provides a complete proven emotional set, or establish comprehensively how different emotional states affect human behaviour and social interactions.

Affective computing has provided several clues by permitting to recognize, interpret, process, simulate human emotions through the different emotional models and their underlying subsystems such as personality, moods, emotional states, etc., with virtual characters. In this paper we will propose a critical evaluation of the most important emotion models used in affective computing and of the relevant subsystems that are required to create believable emotional virtual characters.

Keywords: Emotions, Expressions, Personality, Animation, Speech, Gaze, Facial Expressions, Affective Computing

I. INTRODUCTION

Emotions have always been subject of extensive study, from ancient Greece to modern times, but we are yet to agree on a emotion model that is capable of explaining how emotions react to existent subsystems in our mind such as, personality, mood, expressions, etc. [1]. The diversity and number of emotional models approaches, provide a fundamental clue to how hard it is to create new models or to agree on the role of emotions within the human mind, how they affect our behaviour and social interactions. Within the several existent theories we can distinguish a few that became widely accepted within the affective computing field; Ekman et al.’s universal emotional set [2, 3], Plutchik’s emotion wheel [4], Sloman et al. [5] proposed a model based on primary, secondary and tertiary emotions, and Ortony et al. proposed the (OCC model) the cognitive structure of emotions. There are several differences among these theories, for example; Ekman et al. [2, 3, 6] defined six basic universal emotions, but several researchers [1, 7] argue that disgust and surprise, are too simple cognitively to be considered emotions.

Affective computing is a relatively new branch of computing which aims at investigating and promoting the development of systems and devices that can recognize, interpret, process, and simulate human affects [8].

Recent advances in computer graphics and rendering, permit to achieve photorealistic characters, however the creation of characters capable to recreate human behaviour, emotions, and interact realistically with user or to be perceived as intrinsic part of the scenario and situation, still remains one of the main goals to achieve. According to [9, 10] there are several elements, such as body posture, hand gestures, intonation, moods, facial expressions patterns which are key elements to create believable emotional virtual character.

In this paper we will introduce the most important emotional models in the affective computing field, by discussing the underlying theories, the underlying subsystems and how they interconnect with emotions, or the set of emotional states that compose each model. We will review also, the most important subsystems that are required to create emotional virtual characters and review existent emotional virtual character system.

The remainder of this paper will be organised as follows; in section two we introduce what we consider the most important emotion theories in the affective computing field. In section three we describe the important subsystems required to create believable virtual characters. In section four we describe some of the most important systems from the literature that attempt to create believable emotional virtual characters, in section five we summarise and discuss the issues with existing theories, models and systems created for emotional virtual character. In section six we present what we consider to be necessary future steps for this area, in order to create believable emotional virtual humans. Finally in section seven we present our conclusions and sum-up our findings.

II. EMOTIONAL THEORIES

Extensive research has been undertaken throughout history, about emotions, their origins, what entails them and how they affect subsystems such as; personality, moods, expressions, body posture, etc. They all have been the focus of several theories[1]. We will focus in this section on emotion models which have been developed and implemented within affective computing research.

Plutchik [4] defined a circumplex model of emotions composed of primary, secondary and tertiary emotion pairs, which are combined to create hypothetical constructs. This organisation permits to describe emotions through the use of opposites or similarities with various degrees of intensity. The model defines 8 primary emotions states: anger, fear, sadness, disgust, surprise, anticipation, trust and joy. The model allows blends of different emotion states to different emotional intensities.
Usually this model is represented using a two or three dimensional model, composed by an emotional space in a shape of disk, with different colours representing varying intensity, or by an inverse cone, where the vertical dimension represents the emotional intensity; the base usually refers to the highest intensity in the emotional states. If the three-dimensional model is used, it can be converted into a two-dimensional, and empty spaces in the wheel are considered to be dyads (blends between two emotions). Several systems used this model, such as: Abrecht et al. [11], Kurlander et al. [12], among others [10].

Ekman et al. [2, 3] inspired by Darwin approach, interpret emotions to be an intrinsic part of human evolution and societal development. During his research he concluded that six emotional states deemed to be elementary and universal across different cultures. The emotional states that compose this set are: happiness, fear, sadness, anger, surprise and disgust. However some researchers consider the latter two states too simple to be deemed as emotions [1, 7]. This set also excludes emotions which involves a high cognitive process, such as jealousy and envy considered by other theories [1].

Ekman’s approach due to its simplicity, reduced emotional state set, universality and due to the intrinsic connection between each emotional state and facial expressions as described in FACS [3]. Examples of such systems include Bui’s [13], Greta [14-16]. However Greta also utilises Ortony’s emotional set [7]. Ekman’s emotional set has also became the part of the MPEG-4 facial animation (FA) specification [17].

Ortony et al. [7] created a cognitive appraisal theory of emotions OCC model (which stands for the names of the authors Ortony, Clore, & Collins [7]), which considers emotions to be “valenced reactions" and perceived differently based on the agent goals. This model proposes a structure of the eliciting conditions of emotions and the variables that affect their intensities, rather than observing what emotions elicited the reactions.

This approach is particularly useful in affective computing since it permits to adjust the emotional behaviour of an agent to their goals, standards and attitudes before the actual emotional state is reached. Opposite to Ekman, OCC model recognises that emotions must “always involve some degree of cognition", conscious or not. Therefore, emotions are determined by structure, content, organisation of knowledge representations and the processes that operate on them. OCC does not attempt to connect facial or body expressions to emotions created by such system in contrast to Ekman with FACS [3].

Initially OCC was capable of triggering 22 emotions states but was deemed too complex to model virtual humans [18]. Ortony [10] revised it later to contain only 10 containers, by removing branches that concern other agents.

This approach is particularly useful in multi-agent environments or, in environments where the agent (virtual-human) has to take into account the user emotional state [10].

Aaron Sloman [5] proposed an approach which perceives emotions as subsystems in a multi-layered architecture with primary, secondary and tertiary emotions. Primary emotions are those rooted in very old biological mechanisms, such as startle mechanisms. Secondary and tertiary emotions mechanisms are associated with higher cognitive processes and semantic mechanisms usually which are exclusive to humans.

Instead of perceiving emotions as an isolated system, this model identifies them instead as an intrinsic part of other subsystems in the architecture. However, Sloman recognises that for an emotional system to be recreated it is necessary to obtain, full knowledge of the design and processing architectures. Subsystems are perceived to be different from agent to agent. For example a baby will perceive things differently from an adult or from someone with Alzheimer’s disease.

Most of the emotion theories described in this section, do not focus on expressions. Ekman et al [2, 3, 6] pioneered the study of facial expressions, by mapping emotions and facial expressions together. Facial action coding system (FACS) was created to map the different the positions of facial organs to expressions and consequently their emotional meaning. Most virtual character systems nowadays use somehow Ekman’s [2, 3, 6] work to create facial expressions.

In the next section we will describe how these emotional systems have been used to create virtual emotional characters.

III. EMOTIONS SUBSYSTEMS

While different theorists approach the study of emotions differently, they mostly agree in the existence of subsystems, such as moods, personality, gaze, expressions, language (for example [7] perceives terrified or worried as different intensities of the emotion type fear), cognitive processes, social-culture aspects (according to [2] people in Asian cultures express themselves differently from those western cultures, which may cause somehow difficulties for a visitor into an Asian country to understand certain emotions, state of mind, etc.), etc.. These affect what we understand to be an emotion, what causes an emotion to take place, its intensity and how we express it.

It is impossible to create virtual characters that can be perceived to behave exactly as humans, without tackling each of these systems and the way they interact with each other. However there is little or no agreement to the way each subsystem interacts with each other or, to their exact role within emotions and other cognitive processes. Sloman [5] refers the process of enumerating and understanding these subsystems to be similar to reverse engineering.

In this section we detail and enumerate subsystems which we consider fundamental to the creation of virtual characters.
as believable as possible within current hardware and model constraints.

We can consider personality to be one of the most important subsystems to an agent. It affects the way an agent perceives its surrounding, its social interactions, it also affects the intensity and decay of the emotional states.

Moffit [19] takes a different stand from classic personality theories such as Freud, Skinner and Maslow to create an approach that is suitable for a computational model. He defines personality as “consistent reactive bias within the fringe of functionality”, where personality remains constant during events, whilst emotions are inconsistent and event driven. Several other personality models can be considered, such as the Five Factor Model [20] used in Kshirsagar et al. [16]. It defines five possible personality predispositions to agents, which affect differently the way they perceive social aspects, moods or emotional states. Greta [14] utilizes a personality model based on beliefs and goals architecture (BDN architecture) [21].

According to [10] mood “represents the overall view of an individual’s internal state”. [6, 10] distinguish mood and emotions in three criteria: temporal, expression and cause. Mood is intrinsically connected with emotions but has a longer decay than the latter, and can affect the emotions state intensity. At the time of writing there is no known way to determine how moods and emotions are connected, neither how much they affect each other. Greta [22], Kshirsagar et al. [16] are systems that incorporate mood as a subsystem.

Conversation is one of the most important forms of communication, which together with other subsystems give clues to the agent personality, behaviour and semantic meaning of the messages exchanged. Cassel et al. [23] assume three basic classes on nonverbal communication during speech or conversation, these are:

- Syntactic speech, which depend on the grammatical aspects.
- Semantic functions, which focus on the meaning the speech.
- Interactional functions, which regulate the conversational flow.

During a conversation we can distinguish two main types of behaviour; Speaker behaviour, is mainly semantic [10], its role is mainly to express a message. And Listener behaviour is considered as important as of the speaker [10], its role is to give clues and feedback about the semantic meaning of the conversation.

Gaze is described by researchers as a key social mechanism, used for analysing other individuals [24]. Jellison et al. [25] found that gaze is used in several situations such as dominance, seeking approval or to achieve social intimacy. Studies also show that gaze increases if two individuals engage in frontal conversation, or confrontation, when compared if they are positioned side-by-side. Several models attempted to simulate gaze in virtual characters, Deng et al. [26] proposes a data-driven model that analyses eye saccades signals and maps them to a 1D texture, Fukayama [27] utilizes a two-state Markov Model, where the first state is a gazing state, the second state is an averted state, the model calculates the probability of transition between each state. According to [10] there are no models that detail the relationship between gaze behaviour to their affect in virtual humans.

Facial expressions provide elementary clues to the current emotion state in an agent. Ekman et al. FACS [3, 6], describes parametrically a set of expressions deemed universal across humans. He catalogued these expressions by positioning the different facial parts by devising their position using facial action units. However, Ekman does not consider expressions that are specific to a culture, or even between different humans (for example people that suffered facial surgery), according to [10] these can reach around twenty thousand different facial expressions. Ekman also considers expressions to reflect directly mental states, Carol et al. [?] argue that the context in which expressions happen also affect the way they are expressed.

We consider the subsystems above to be the most used and the most important ones when creating believable emotional characters, in the next section we will review some of the most important virtual character systems, in which some or all of these subsystems are implemented.

IV. EMOTIONAL VIRTUAL CHARACTER SYSTEMS.

In the past fifteen years, several systems have been created by researchers or industry to attempt to simulate emotional virtual human characters, some achieved great results, which provided a benchmark to other systems in affective computing and virtual character animation. The range of the subsystems implemented in each of these systems varies, but so does performance and realism which are intrinsically related.

There are several possible ways in which we could organize these systems, but since we are targeting emotional aspects in these systems, we will consider two categories described in [10], static and dynamic emotional systems. Below we present a selection of the most important contributions to this research area:

- Kshirsagar et al. [16, 28, 29] proposes a dynamic emotional system composed of three main layers, composed by: personality, mood, and emotional states. The Personality subsystem is the top most state, it follows personality states from FFM [20], where each state is composed by a Bayesian Belief Network. Personality remains mostly unchanged during the timeline, but it causes deliberative reactions that affect moods and consequently emotional states. Moods states were simplified to match OCC emotional categories, with three states; good, bad and neutral, which are affected by the agent personality and the emotional state. While mood switching is not a frequent task, the frequency in which it happens is dependent on the personality model and emotional tags received through text from the user. Emotional states are composed of 24 emotions, 22 emotions defined in OCC model [7], plus
disgust and surprise which belong to Ekman’s [2, 6]. Ekman’s expression set [3, 6] are mapped to match the 24 emotions, an emotional state is decided upon three factors; from user input, current mood, and previous emotional state. These states are handled with probability transition matrices for each mood. Once an emotional state is designated it is mapped to a facial expression and synchronized with speech.

- Bui [13], created the Obie architecture, a dynamic emotional model, which creates an “expression model selection”, based on Ekman’s basic emotion states. Two fuzzy rule-based systems map an emotional state vector, provided by an emotional model. One rule maps a single emotion to its universally recognized facial expression equivalent, the other rule maps two emotions with the highest intensities to create a blend that represent intermediate facial expressions from Ekman’s set. Bui created also an “expression mode selection” which determines which emotion should be displayed, or if one or more expressions are necessary to be displayed using a set of rules. However, humans can generate around 50,000 distinct facial expressions [10], which this emotion model cannot guarantee that the emotions created will belong to this extensive set. Neither can it guarantee that these are represented realistically or that the user interprets the created expressions in accordance to what they attempt to represent.

- Irene et al. [11] proposes a static based emotional physics-based facial animation system, capable of simulating emotions and facial expressions. It model expressions using Plutchik’s emotion wheel [4] to Cowie et al. [30] emotion disc, which distinguishes emotions intensity based on their distance from the disk centre. The centre represents neutral emotional states and the extremities represent emotional states with the highest intensity. Facial expressions can be generated using two different methods, if the target facial expression is part of an existent emotion state where only intensity changes, where it attempts to map the resulting facial expression to the result. If a target facial expression does not fit within a specific emotion, it maps the two closest emotions and interpolates between the facial expressions associated with these.

- Greta [14, 15] is another complete dynamic emotional system which implements the MPEG-4 standard for virtual character animation [17], its mind is modelled according to beliefs and goals architecture (BDN architecture) [21]. The use of this system permits subtle changes in the agent mind and in the emotional intensity during time. The connection of dialog meaning to beliefs and goals, it defined four communicative functions; belief, intention, affective state and metacognitive information. The personality system implemented in this system allows changes through time, to match changes in the agent goals, which will affect the agent emotional decay. Greta disguises the time that takes to analyse the connection of goals, intents, to the meaning in user dialog by utilizing gaze and facial movements to simulate “thinking activity process”, “remembering a fact” or trying to “make inferences” [14]. Greta defines its emotional state using emotions from Ekman’s universal emotions [2, 6], a subset from OCC [7], and from Keltner [31] including embarrassment and shame to enhance the limitations of Ekman’s [2, 6] and Ortony [7] emotional sets.

Multiple emotions can be modelled during a timeline, what permits new emotions to be created while others decay. BDN’s [21] are used as a monitoring system that calculates the probability of achieving a determined goal, to map it to the agent personality and produces the result emotions and time decay.

The systems described above usually have their advantages and drawbacks. Static emotional models attempt to determine the correct emotional states, they do not account for emotional states where there is a degree of uncertainty, neither attempt to dynamically create new expressions or understand their meaning. Static systems do not account for expressions particular to a culture or the ones that are specific to individuals, for example; where a person is subject to surgical intervention in the face or body, twitches that we acquire which represent states of mind or particular emotional states, per example; some people twitch when they are lying or feel guilty or nervous about something. Systems that implement these static models usually suffer from lack of consistency, since the use of static emotion representation enables to display anger consecutively to a happiness expression [10] without considering intermediate emotion states or time decay. Dynamic Emotional systems attempt to answer to the issues caused in static emotional systems, but the relation between time decay, expression blending or probabilistic calculation of emotional state is not completely understood, therefore they cannot guarantee to achieve always a human like emotional behaviour. Furthermore, usually such systems are not suitable for real-time use. In the next section we will summarize several existent issues in affective computing and surrounding fields that study emotions.

V. CURRENT ISSUES

The lack of consensus on the actual emotion definition or on what constitutes an emotional state summarizes the uncertainty that exists in the study of emotions. While there is such degree of uncertainty we cannot create virtual characters that behave and act like humans.

Current emotional theories, offer either a very simplistic approach which ignores the cognitive importance of emotions allowing only to create what Sloman [5] considers a shallow emotional theory, or offer an ambiguous approach [32] or are too complex such as Sloman [5] and Ortony [7].

Ekman [2, 3, 6] attempts to find the most elementary emotions, however according to [1] surprise and disgust are not considered emotions, due the lack of cognitive process associated with them. Ekman’s emotional set is also considered too simplistic to simulate virtual characters. Usually emotional simulation systems such as [3, 16]
complemented this set with OCC [7] emotional states. Ekman does not consider emotional decay, their intensities or emotional dyads, or multiple emotional states to exist in an agent. Ortony et al. [7] proposes a complex system which has been devised logically ambiguous by Steunebrink et al. [32]. It also has been found too complex to be integrated within a virtual human [18].

Sloman [5] proposes a model that is too complex to be created without a complete understanding of all surrounding emotion subsystems. However we are far from reaching full understanding of each subsystem, and how they relate and affect each other. For example little is known on how personality or mood affects an emotional state. Most research into facial expressions only consider expressions present from Ekman universal emotions [2, 6] and FACS [3]. If a system decides to implement a more rich palette of expressions, there is little knowledge on what facial expressions reflect emotions such as shame, annoyance, etc., which are not present in Ekman’s universal emotion set [2, 3, 6]. While Bui [13] attempted to address this issue by generating dynamically expressions, it fails to recognize that some expressions are local, and that different cultures may express differently. Bui [13] does not give guarantees that the expressions generated are actual expressions, or that they will be recognized by the user. Finally most implemented approaches perceive these subsystems as single units with little or no interaction, when in fact the existent interconnection is much deeper, affecting each other at all times.

The issues discussed above account in our view to be the most important in this research area, in the next section we will attempt to provide what we think will be possible future directions to the affective computing field.

VI. FUTURE DIRECTIONS

Affective computing became a very active research area in the past forty years which resulted in several emotion theories, with very different approaches. The larger number of research projects in this area, permitted several advances to occur but more importantly permitted to understand how little we know about such intangible field. Therefore it is extremely difficult to point with certainty to which areas would research produce better results, or where we are closer to achieve definitive results.

As Sloman [5] correctly stated, without knowing how each subsystem works precisely it is impossible to recreate human behaviour in an emotional system, therefore as this research area evolves it will be necessary to create a standard with a concrete specification in affective computing to permit quicker research progress. If such specification is created it would permit quicker evolution within, other related disciplines such as; psychology, philosophy, neurobiology, and communication. It is also necessary to agree on the emotion definition, what is considered to be emotional states, how cognitive processes affect or are affected by these states.

Recently several studies attempted to create means to recognize and catalogue expressions into databases [33], however most of these studies focus on expressions recognized in FACS, while systems such as Greta recognize [14] the need to expand Ekman’s set, it is necessary to recognize a final set of emotion states, to map precisely facial and body expressions to these emotions, in order to achieve a general emotional and expressionset. While systems such as Bui [13] attempt to dynamically select and create facial expressions, there no knowledge between the relation in opposites expressions such as, happiness or anger and how they are affected by previous emotional states, it is then necessary to research and evaluate the relation between these.

Our contribution to the affective computing area will focus in integrating within our framework Charisma [34] a way to observe and establish a relationship between each emotions and the expressions that can be represented by emotional state. We will also study the relationship between each expression in an emotional state to an opposite emotional state. Our research focus will focus on the organization, cataloguing and on the visual and semantic relationship between each emotional state, and expression associated with it. We believe that these approach permits to catalogue emotions states and expressions while attempting to find relations between these.

VII. CONCLUSION

In this paper we attempted to provide an in-depth review of what we consider the most important emotion theory models in affective computing field. We described what we consider the most relevant subsystems in emotional systems, and provided an insight into several applications that attempted the creation of emotional virtual humans. During this review we concluded that several advances have been made in this field but, much more theoretical as well as applied research is needed into human behaviour, to recreate human emotions in virtual characters. However the current advances in this area permit to point at several subsystems which we think are necessary to consider when attempting the creation of an emotional virtual human character. The acceptance of virtual character by the user largely depends on its behaviour and capacity to communicate with its surroundings or the user itself in a believable fashion. While verbal communication and expressions are very important to make a character believable, it is the non-verbal behaviour that will make a character to be truly accepted by users. While there is no certainty to the correct length of time that an emotion must demonstrate itself, we do know that an emotion state is short lived when compared, with a mood state or a personality state, therefore it is necessary to account to these differences when designing an emotional character. If a virtual character demonstrates an emotion, it shall change the intensity to how it feels it based on personality, mood, and preferences to its surroundings as Ortony [7] and Sloman [5] well refer in their theories. When transitioning from different emotional states, it is necessary to update the subsystems according to the new emotional state and prepare a realistic emotional decay of the current emotional state.

In fact, an implemented system is only as effective as each one of the parts that compose it. To create an emotional system that attempts to be believable by a user, each
subsystem will have to be capable of: interconnection, responsiveness, awareness of the surroundings and situation. Until a system with such key aspects is created, its acceptance and believability by a user will be limited.

VIII. REFERENCES