The Chatbot Who Loved Me

Gábor Tatai
Department of Computer Science
University College London
London, UK
g.tatai@cs.ucl.ac.uk

Annamária Csordás, Árpád Kiss
AITIA Inc.
Budapest, Hungary
{acsordas;kysss}@aitia.ai

László Laufer
Department of Psychology
ELTE University of Sciences,
Budapest, Hungary
notme@izabell.elte.hu

Attila Szaló
Department of Measurement and
Information Systems
Technical University of Budapest, Hungary
aszalo@mit.bme.hu

Abstract

The paper describes the BotCom complex embedded conversational agent system, which is accessible over the Internet and capable of high quality natural language processing as well as sophisticated manipulation of emotions. It relies on an extension of Plutchik's emotional model. Experimental cases leading to useful constraints for optimization are discussed. BotCom, a prototype for a robust commercial application, is the outcome of both research and realistic development.

1 Introduction

Research on embodied conversational agents (ECAs) has intensified in the past few years as avatars representing users providing virtual presence over the network as well as autonomous chatterbots have become widely applied by both the research community and industry. The most apparent application areas of ECAs are education, entertainment, search engines and general tools for information retrieval, virtual marketing, customer care and support etc.

It seems that the number of questions about theoretical and practical implementations are growing exponentially as we head towards more and more complex applications. Some of these questions seem to be really tough both in theory and practice: application of emotions, dialogue management, recovery of semantic information, and the synchronization of all these. There are several papers addressing these issues and suggest useful partial or global solutions (Cassell and Vilhjalmsson, 1999).

There is no space to discuss all the details in this paper of using ECAs in human-computer interactions. Therefore we focus on some highlights where, we think, our ECA is special or where a theoretical or practical observation proved to be particularly useful, in the hope that others might benefit from these as well.

2 System Architecture and Overview

2.1 BotCom

"BotCom" is an existing prototype system, part of a research project and also currently under introduction to industrial use. It is capable of chatting with users about different topics as well as displaying synchronized affective feedback based on a complex emotional state generator, GALA.

Moreover, it has a feature of connecting to various information sources and search engines thus enabling an easily scalable knowledge base. As the operation area and access medium of BotCom is the Internet, all constraints that apply to this medium affected the possible implementation scenarios of the system. The primary use of BotCom will be interactive website navigation, entertainment, marketing and education.

2.2 Language

The language of the ECA is Hungarian, but with adequate language and semantic (WordNet) plug-ins the default language can be changed. There are several language dependent specialities, for example Hungarian is an agglutinative language. During the sentence processing we not only have to cope with hundreds of different appendages and their variations, but in some cases the stem of the word changes as well. The different appendices added to the end of the words can change the meaning completely. The same words with different punctuations can have totally different meanings. We also have to take into account that users may mistype words, leading to misunderstandings. All these make it really challenging to create an ECA or even a simple chatterbot in this language.

2.3 The Middle Way

In developing BotCom, we have faced several competing demands, e.g. to raise the complexity of interactivity of the ECA chat capabilities, and to minimize the resources needed during usage. To arrive at an adequate complexity of the pattern matching subsystem for the emotional model, the facial expressions, the number of gags (small sequence of funny animation) and gestures we use, we conducted several experiments. Our goal was to find a way not to overcomplicate or oversimplify the system, but something between in order to satisfy our most important priority: to create a workable system. Therefore, our working methodology was to find an equilibrium among these constraints, i.e. a *middle way*, with the corollary that more is not necessarily better.

2.4 Architecture and Technical Considerations

Let's take a glance over the system architecture in order to understand better the functionalities of BotCom. It is primarily a client-server architecture with a separate chatterbot development module, BotMaker (Fig. 1.).

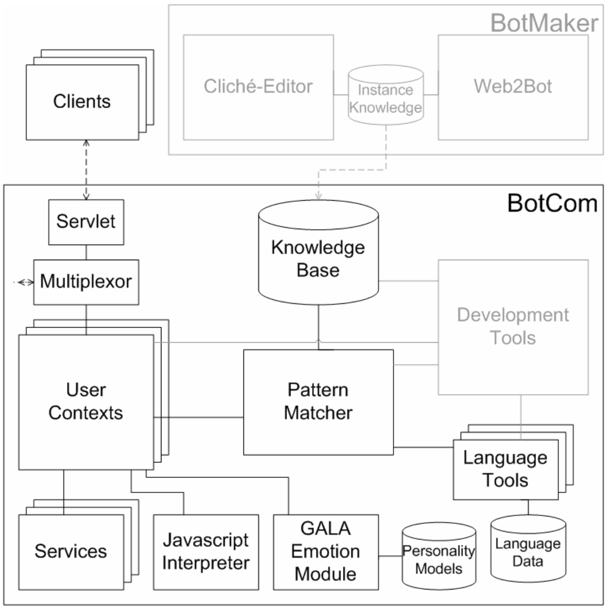

Figure 1. The main components of BotCom and the BotMaker service and development tools.

Clients

The ECA is displayed in a web browser. The client visualizes the messages of the chatbot, the animations representing its emotions and the gags, as well as forwarding users' messages to the server.

The web pages where the ECA is embedded into are created and maintained by the Webra portal management tool (also our development, an independent Linux/ Apache/PHP based tool). Integrating the ECA into a portal engine is advantageous, because the knowledge base of the bot can be extended by expressions from the web site, since the portal engine stores the keywords and full articles in an indexed database. Thus the integration enables the access to this database.

Animations

Animations are played by a Flash movie clip player. We use pre-rendered 3D bitmap animations for all basic emotions, entertainment and personalized gestures.

There is an important tradeoff here that requires some explanation. Many ECAs use eye/lip synchronization along with natural language processing. We have evaluated many of them according to recent guidelines (Buisine et al., 2002; Isbister and Doyle, 2002; Ruttkay et al., 2002; Xiao et al., 2002). Those ECAs use either 2D animations or 2D Flash animations in vector format. But the 2D animations are not so attractive according to our surveys, not only because of 2D, but also the tiny size of the faces. There are also experiments that use 3D models with the VRML 2.0 player, that recognizes the body description language HANIM. However, to play a VRML animation on the client side, one needs to download not only the animation data, but also to install a VRML plug-in for the browser which is resource-consuming.

As an alternative, therefore, we have chosen to use pre-rendered animation segments.

Server

On the server side a Java servlet collects the messages and transmits them to a multiplexer that queues or distributes them among the servers for parallel operation. Messages are processed separately, in the appropriate user context. Every logged-in user has exactly one user context, which is responsible for the communication and stores all information about the user. The user messages are handled by the dialogue system which matches them against the patterns stored in the knowledge base. The answer, which consists of the status obtained from the GALA emotional model and the browser controlling commands (if any), is sent back to the client afterwards.

Infoserver and Agentcities

The BotCom is connected to an external information source server, the Infoserver. This application is capable of collecting data from various sources on the Internet and can be used to provide dynamic, up-to-date information to the Chatbot, such as weather reports, stock exchange quotes and other useful data. Probably the most important part of the infrastructure from a future perspective is its connection capability to the AgentCities EU Network project. It aims to connect cities mostly across Europe into a service pool of agents, providing up-to-date travel, library, weather etc. information on a standardized platform. The network standard is elaborated by the FIPA foundation (www.fipa.org) and the project is heavily funded by the EU. BotCom understands FIPA compliant ACL (Agent Communication Language) messages and is able to connect to the Hungarian backbone of this service located at the Computer and Automation Research Center of the Hungarian Academy of Sciences. Via this service it is able to communicate with other agents currently available over the platform.

Development tools

Development tools for uploading and managing the knowledge base of the ECA and adding functionalities include the BotMaker and the Cliché Editor.

3 The GALA Emotion Module

The most distinctive part of the system is the emotional modeling and processing module. It has two main parts, a runtime and a development module. The former acts as a black box within the system, processing messages and converting them into emotions. The latter is a visualization and experimenting tool using the Java 3D API. We test the applied emotion model and the dynamic behavior

Figure 2. The Interface of the GALA synthetic emotional modeling, generation and verification system

of emotions with this interface. In the following section we discuss GALA in detail.

3.1 Emotional Modeling

Motto: "Everyone knows what an emotion is, until asked to give a definition." (Fehr and Russell, 1984)

Modern chatterbot systems do not really pay ample attention to emotional modeling or the solution is far from being sophisticated. For a human being a communication partner without even the simplest emotions can be very boring. Not mentioning tutoring and general educational systems which are one of the primary potential application areas of the chatterbots. Probably most of us have had unpleasant experiences with poker-faced teachers at oral examinations, when there was no apparent sign of understanding, approval or disapproval. Therefore it is almost a mandatory requirement to equip an ECA with a reasonably refined emotional model. Such a communication partner can be more alive, convincing and entertaining as noted already in earlier experiments (Rizzo, 2000).

Thus we aim to develop a more sophisticated emotional module, enabling the treatment of the continuously changing emotional status during a conversation.

Our system, the GALA Hierarchical Integrated Synthetic Emotional Modeling, Generation and Verification System has the following multiple functions:

- **Modeling:** The module is able to simulate the emotional status of the chatterbot and the affective content of the conversation elements.
- **Analysis:** Enables continuous investigation of the messages and the emotions as well constantly monitoring the changes.
- **Synthesis:** Generates the actual emotional status and the emotional behavior based on the emotional affect of the messages.
- **Personalization:** Enables the creation and management of various chatbot personalities – each can behave differently in the same emotional situation according to their configuration.
- **Testing and evaluation:** Helps calibrating the connection between the messages and emotions and the verification of the basic emotions combined into more complex emotional structures.

3.2 Affective Computing

Affective Computing is a flourishing research area. According to a definition by a leading researcher in affective computing, this field deals with problems where "computing is related to, arises from or deliberately influences emotions" (Picard, 1997).

Various recent advances in psychology and neurology are making cognitive scientists and psychologists to rethink the origins and behaviors of various emotions. Therefore there is presently no universally accepted emotional model.

On the other hand the early work on emotions by Darwin and others who distinguished discrete categories, so-called basic emotions (Darwin, 1872; Tomkins, 1962) have resisted the test of time at least for their main categories and the description of these emotions. Others, however, have emphasized the continuous dimension of emotions (Schlosberg, 1954). The continuous approach is likely to be closer to the real operation of emotions. Nevertheless, it seems that the *reflections* on the emotions which identifies them within the self and thus makes the self aware of them, introduces a discrete element into the anyway continuous neurobiological system. The discreteness is even more noticeable when the emotion is translated to language and expressed so (Oatley and Jenkins, 1996).

The eight basic emotions described by Tomkins seem to be the most widely accepted (fear, anger, anguish, joy, disgust, surprise, interest and shame). This set is a reasonable extension of Darwin's first identification of the basic set (joy, anger, fear).

Beyond these, there is a well-known and widely accepted classification of emotions. This is the OCC Cognitive Model (Ortony et al., 1988; Paiva, 2000) which provides not only a definition of them, but also a hierarchy based on the target of the emotion (self, other) and other meta-categories such as consequences of events, aspects and actions. However, this is a strict and complex cognitive model of the key emotions. In order to use this approach, one does not only need to model the emotions themselves, but it is necessary to adopt the entire cognitive model including all its processes and structures. In that case we would be able to manage the emotions appropriately, but it would also be very cumbersome. Even a simple model here implies a complex architecture, as is clearly visible from reports on similar projects (Integrating Models of Personality and Emotions into Lifelike Characters) at DFKI (André et al., 2000).

3.3 The Plutchik Model

Our goal was to select a sophisticated emotional model complex enough to provide ample workspace, but not bound to a net of cognitive processes. We have found that the model suggested by the psychologist Robert Plutchik is an appropriate starting point for further description and generation of synthetic emotions (Plutchik, 1980; 2001). He defined eight basic emotions that are just a little different from those of Tomkins. In addition he suggested another dimension, intensity (with 3 levels). Thus we get 24 emotional states. A novelty of his proposal is that the emotional space is mapped to an upturned 3D cone, where the positioning of a particular emotion reflects psychological distances and intensity differences between states. The origo is in the apex of the cone expressing neutrality ("no emotion"). We have selected this model for implementation and incorporated it as a part of our more complex, layered architecture.

We follow a decomposition approach to explain the GALA model, starting from an example expression and tracing it until the system generates the dominant emotion of the ECA as an output.

4 The GALA Hierarchical Emotion Processing Model

4.1 Layer One – Message Act Processing

In order to create a usable synthetic emotion model for a ECA/chatterbot that communicates principally via text messages, one has to design an appropriate mapping scheme between the emotions and the expressions that the chatbot sends or receives. The connection between the text messages, or better to say, *message acts* and emotions is thoroughly investigated by M. A. Gilbert in his work (Gilbert, 1997; 1999). This work was particularly useful because it has pointed out that there is a reasonable difference between the well-known concept of a *speech act* and a *message act*: the meaning of the latter also carries emotional load. According to Gilbert's definition, a message act is similar to speech acts: "A message act, being analogous to an utterance act involves an expression of emotion that is identifiable to the recipient or observer." (for example: threat, fear etc.). The speech act is relevant for analyzing sentences, querying semantic information and emphasizing the verbal aspect of the communication. In contrast the message act focuses on various forms of the communication and is used to express dialogues and emotional aspects in addition to the semantic (meaning) level.

When processing the input expressions, initially they are passed to the GALA processing module. In the first step the conversational elements (words, expressions, sentences) in the dialogue are assigned to the emotional message acts which are processed at the topmost layer. If GALA finds a match of the input expression it passes to the second layer, otherwise the element is dropped and does not alter the emotions of the chatbot. The set of recognizable elements can be extended and the system can be trained. The method of training is described later.

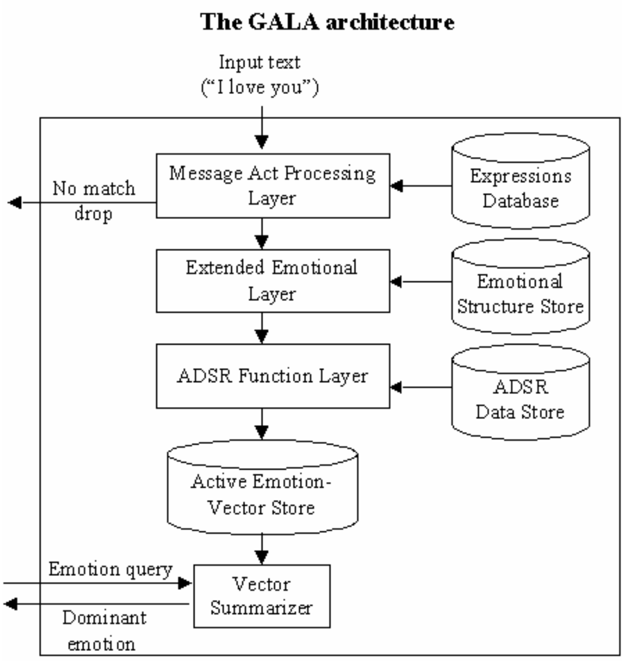

Figure 3. The architecture of the GALA engine and the processing of a message act.

4.2 Layer Two – Extended Emotional Layer

A message act almost never expresses a pure basic emotion. Therefore we have defined a second layer which allows the construction of mixed emotions that are more closely related to message acts. In this layer one or more basic emotion components can be assigned to the individual message acts. These components can be weighted, so a message act could be represented by a mixture of e.g. 20 % sadness, 40 % anger, 40 % fear. There is a predefined database, the "Emotional Structure Store" (ESS) which allows the storage of these complex emotional structures and their assignment to the appropriate message acts. The basic emotion components are represented by vectors and thus the resulting mixed emotion is also stored and manipulated in a vector. On the 3D surface suggested by Plutchik, the origin of these component vectors is the apex of the cone, and they point towards the center of the particular surface segments that are assigned to that basic emotion. In general we did not introduce a limitation on the number of basic emotions that can build up the resultant emotion which is associated with the message act. However, according to our experiments and anticipated real-life requirements, it seems that there is a practical upper limit. It appears that most of the expressions can be realistically covered by a mixture of at most 3 or 4 of the 24 predefined emotions in a given situation. The screenshot of the emotion editor (Figure 2.) gives an illustration of the behavior of the emotion vectors.

In summary, the function of the second layer is to break up the message acts into basic emotions taken from the Emotion Structure Store and forward them to the third layer.

4.3 Layer Three – ADSR Refinement Function

It is well known that emotions are time-dependent, so we had to manage that the amplitude of their synthetic pairs change over time, as this is not in the original model. In addition, the characteristics of this change vary by emotion. We do not need to imitate chemical or neural aspects (which are anyway unclear) of emotional change, so we have selected an approach that is more controllable and probably close enough in its effective behavior.

Therefore, in the third layer we have used a well-known function, ADSR, to describe the change of emotions over time. Signal processing and musical applications widely apply the method of dividing a signal into four main parts. Thus an envelope of a signal can be defined by the Attack, Decay, Sustain and Release values.

In our model all basic emotions can have their own ADSR functions, which ensure unique time characteristics for each. Moreover, the intensity of each emotion can be tuned in order to provide scope for relative scaling of the emotions. So, in a particular situation the amplitude (intensity) of joy can be stronger than the intensity of the rest of the emotions while still maintaining the same ADSR function and thus its behavior over time.

4.4 Computing of the Dominant Emotion

The ultimate goal of the process is to determine the dominant emotion since we have a strong constraint – we can play only one animation at a time. The third layer, after the decomposition of the *message acts* into emotional structures and further into basic emotions, selects the ADSR and intensity values corresponding to time t ("now") for each basic participating emotion. These emotion vectors are kept in an Active Emotion-Vector Store (AEVS). The actual *resulting emotion vector* is calculated by summing and normalizing these active vectors. This final vector contains the effect of the actual message act (through its decomposed participating emotions) and the fading effect of previous message acts which are already in the Active Emotion-Vector Store.

The processing of the normalized sum of the vectors produced by the three layers can be expressed as follows. Please note that we do not use simply the formula below: we have a pipeline process where results from previous time instants are also taken into account. The formula is provided for elegance and better understanding only.

$$\underline{r}(t) = \frac{\sum_{k=1}^{n} \underline{e}_k(t)}{n} = \frac{\sum_{k=1}^{n} i_k \underline{E}_k f_k(t - t_{0k})}{n}$$

- $\underline{r}$: resulting emotional state vector
- t : actual time
- n : number of active emotion vectors (where "active" indicates a non-zero vector)
- $\underline{e_k}$: the kth active emotion vector
- i_k : the maximal intensity of $\underline{e_k}$
- $\underline{E_k}$: unit vector of $\underline{e_k}$
- f_k : ADSR function belonging to $\underline{e_k}$
- t_{0k} : the activation time of the kth emotion

The result vector will point to some location within the cone. The *dominant emotion* for that *t* moment will be that basic emotion (out of 24) which is the closest to the endpoint of the vector sum.

At successive times, the system dynamically recalculates the emotions by repeating this decomposition process from layer one to three. Therefore a nice dynamics and pattern will have displayed as emotions appear and fade. This process can be traced with the emotion editor which not only includes three rotateable and zoomable camera views of the cone with the vectors inside, but the basic emotions diagram too. This diagram indicates the percentage shares of the basic emotions from the current resulting emotion vector (see Fig. 2.).

4.5 Extension of the Second Layer

When implementing this model, we have found that further refinement is necessary when assigning the ADSR parameters. Therefore, we have extended the possibility of function assignment to the second layer. In practice, when a set of emotions is stored in the Emotional Structure Store (ESS), each basic component is stored and then all are used together. Each of these basic emotion components can have a separate ADSR function, which overrides the general basic emotion ADSR if the corresponding ESS structure is triggered by the appropriate message act.

This feature enables more refined tuning of emotions and, as a result, a more natural fitting to the real effect of a message act.

For instance, if the "admiration" basic emotion is combined into a more complex emotion structure that describes "coaxing", the time-dependent behavior of "admiration" can be very different (e.g. shorter) from what it is within the "submission" ESS structure. The latter can consist of the "terror" and "admiration" basic emotions, for example, according to Plutchik, where "admiration" can be a more lasting component in the structure, depending on the context.

One may notice that our original standpoint was to create a fairly straightforward emotional model. By now, however, it seems to be moderately complicated. In fact, most of the components (such as the emotion handling and the ADSR functions) are quite simple and reused. The only thing that happened is that we have used them in multiple layers thus enabling complexity with simple, reused and override-able components. This approach behaves similarly to the theoretical and experimental results of psychology, where complex memory and emotional models arise from using very simple elements.

5 Evaluation of GALA

5.1 Testing the GALA Modeler

When evaluating the emotional engine we have faced several interesting issues. Many of the results met our expectations; however, some non-trivial questions have arisen.

In order to use a chatbot system equipped with emotional capabilities it is necessary to create a database that contains the mapping of words and expressions into emotions and more complex emotional structures. We have found the idea itself and the creation process for this mapping database to be analogous to speech and language databases that one has to create when working with speech recognition or translation systems. The approach to all these databases is to take a set of linguistic or phonetic components and train the system both manually and automatically.

We had to calibrate the emotional engine in order to map the message acts into emotions. The visualization module of GALA is capable of reading short texts, such as lyrics and poems. We tried to choose texts where the expressions of the emotions were not hidden or metaphoric. Obviously, the handling of deep semantic information, multiple, metaphoric or analogous meanings are beyond the capabilities of current systems in general.

The user has a possibility to label parts of the text (words, expressions or entire sentences) with message acts or basic emotions. Various colors correspond to different labels. The emotional engine can process the text from the beginning and calculate, track and display the change of the internal emotional state while reading. When we were observing a trace of a poem, it was apparent, that the pace of change was very fast compared to real-life emotions during reading.

Since the system relies on text input only, it has to be much more sensitive to its information source. It is frequently said that "my emotions are in a whirl". Let's call this feeling analogous to the processing of some particular (short text) poem. The crystallization of the final emotion then could be analogous to the final emotional state that we get into as a result of the analysis (reading) of an entire lyric or poem. If we find that the system arrives at the same state after processing the entire text as an ordinary person who reads the same text would reach, then we may conclude that the model has mapped the message acts appropriately into emotions on a macro scale, not considering the micro level, which would involve a word-by-word change in our case.

In these experiments we have monitored how realistic the changes of the emotions were in the artificial personality and whether the cumulative effects of the message acts would lead to an appropriate reaction in time. One key to this is the analysis of possible emotional "phase transitions" in our model.

5.2 Original Transitions

First of all we have to emphasize that the Plutchik model itself allowed easy transitions in the following cases:

- When the emotional intensity has changed, it could move into a similar emotional state with stronger affect (e.g. from joy to ecstasy)
- An emotion could easily change its immediate neighbor. It was Plutchik's original intention as well to place the similar emotions next to each other. The movement on the surface of the cone realizes this transition.
- When an emotion has a low intensity, it could easily switch into another "light" emotion, since their geometrical distance is small (e.g. from acceptance to boredom). On the other hand, transitions between states that are far away geometrically are less likely (e.g. admiration to boredom).
- A similar effect is the switch between very strong emotions. According to the model, a change between strong emotions is less likely than between light emotions. This is also a result of the geometry.
- For the sake of accuracy it is important to note that in Plutchik's model emotions, moods and cognitive states are mixed together somewhat (e.g. boredom is more a cognitive state than an emotion etc.). But since we are building an approximate, synthetic model, we can afford this more unified and "standardized" approach. Also, Plutchik's own view strengthens our position, as he argues for the unified treatment for good reasons. For instance, many of the differences among these emotions and cognitive states are based on nothing much more than time differences. An episodic emotion can turn to a cognitive state if its longer in time etc. This effect is handled by the ADSR functions in our extension of the model (Oatley and Jenkins, 1996; Plutchik, 1992).

5.3 Adequacy of Behaviors

- Most of the basic observations related to the original Plutchik model were relevant and worked well. This included the change of the length of the emotional vector (intensity) that resulted in realistic emotional state transitions (rage-anger-annoyance etc.).
- Effects similar to each other have strengthened the resulting effect; contradictory effects (pensiveness – serenity) weakened or neutralized each other.

5.4 Animation–Reply Discrepancies

Synchronization between the response text and the facial expression is a difficult issue. The emotional model provides a mapping of the discrete facial expressions to the discrete space of emotions. Therefore it will always be possible to assign a preset animation to the resultant emotional status; however, that will not be the case when selecting a corresponding message act. While the facial expression and the detection of complex emotions from the input text are realistic, it is impossible to adapt and prepare the system to contain all types of answers in all topics for all emotions. Since we are working not only with preset animations but with preset answers as well, it would create an enormous database – unnecessarily. Nevertheless, the chatbot has to give something in answer to the input. It tries first to select an emotionally relevant answer, but if this is not possible, it will do something that we cannot always predict or guarantee. There should be a reasonable tradeoff to cover the most frequent emotions and provide replies to them, as well as many neutral replies. In this case even if the reply might contradict the facial expression, the probability of a sharp conflict will be low. We intend to conduct further experiments on this matter of synchronization.

5.5 Vector Calculation Issues

Although the summing operation on the vectors for obtaining a resultant vector seems to be a straightforward and computationally economical solution, by no means does it provide the appropriate behavior trivially. So, we have to ask ourselves, if is it a good idea to "sum up the emotions" into one resultant emotion?

- Of course, in everyday life we cannot talk about the "resultant vector of our emotions". However, when the time comes for a reply, humans also act on the basis of a current dominant emotion and they express it with small variations. Therefore the idea of selecting a dominant emotion is not so far from real life. We were anyway forced to select an unambiguous final emotion since we had to use it for picking the appropriate animation and seek for a relevant text reply. Multiple emotions could be displayed only if we were using a face built up from separately controllable parts, but that was not the goal as we have discussed it earlier.
- The dominant emotion selection method can be based on a "weighted summing operation", which is present in humans in a special, unconscious form. It adds up all emotions that happened over time and provide a store, or imprint. That emotional memory is query-able, and one can retrieve the dominant emotional state for example by asking – how are you? Thus the synthetic modeling process may approach this calculation problem similarly.
- In several cases the sum operation of some emotional effects does not deliver the expected emotional result status (e.g. $\underline{r}_{pensiveness}$ + $\underline{r}_{annoyance}$ =? $\underline{r}_{boredom}$, which is not really what one would expect from a lifelike model). This issue was clearly foreseeable and was caused by the geometrical nature of the emotional space.
- In our calculation the most important emotions have the most significant vector components, therefore they are dominant after the summing as well, hence they strongly influence the resultant vector - which looks like correct behavior.
- The summing provides that if there are many active vectors at a certain time, the resultant will be close to a neutral state. This is close to the natural experience of emotional overload (and chaos), where the mind collapses and in many cases a limp or passive status occurs. It is an advantage. However, in other situations the presence of multiple vectors can result in an irrelevant final emotion. This shortcoming can be corrected by a refinement of the calculation method. One trivial possibility would be to reduce the maximum number of participating vectors to two, since most of the statuses that are hard to interpret appeared when there were more than two significant vectors present.

5.6 Issues with Setting the ADSR Functions

- It is not easy to pre-program the behavior of the ADSR functions. If the sustained duration of the emotions is too long, too many emotional components will be present at a given time, which may result in emotional overload and the ceiling effect, which we discuss shortly.
- If the sustained interval is too short, the effect of the emotion might be taken wrongly as insignificant.

There are possibilities to resolve this conflict. One possible improvement can be the introduction of a relation matrix, where emotions can inhibit or stimulate each other. In this way, if a new emotion occurs, it could block the effect of long-term, previously-arising emotions. There is an analogue of this effect in real life. It is well known that the most recent effect sometimes has a more important role than it should have – humans optimize greatly by using the time as a sorting base.

5.7 The Ceiling Effect - Emotional Overload

The so-called *ceiling effect* occurs in the emotional system when the intensity of a certain emotion reaches the theoretical maximum – in our case geometrically the topmost circle of the cone. When adding a new emotion vector that influences the previous (already at maximum intensity) intensity state, the resultant vector will not modify correctly the overall emotional state. This is an emotional overload status when the predictive ability of the model is lowest. In psychology there is a similar phenomenon that after some given time, when the level of emotions (or the general arousal level) is kept constantly very high, the psyche tires out. As a result the "sensitivity" of the general emotional system decreases. This means that the same stimulus does not have the same effect in this overload state. In theory the model could be modified to accommodate infinite intensities, but it would not be realistic. Therefore the fixed ceiling is a reasonable element of the model, but one should be aware of its limitations when the overload situation occurs.

In the GALA system the possibility to manipulate the shape of emotion functions with the ADSR editor technically allows the agent to modify the characteristics of its own emotions. This could be done with a feedback loop so that when an emotional overdrive occurs, the agent could decrease the amplitude of the ADSR curve corresponding to the overdriven emotion. Thus it could reduce the chance of another overload. This automated feature is a subject to further development.

6 Conclusions and Future Work

We have briefly described the BotCom complex ECA. Although several key features were discussed, we did not have the opportunity to elaborate on the dialogue manager, the knowledge base and its content, example discussions with the system and their analysis. Neither we had the possibility to detail personalization features and how the system remembers the user and exploits this information in further dialogues.

Nevertheless we could share some useful conclusions from the analysis of the emotional model, such as the applicability of the Plutchik model in real-life applications. It seems that when there is a need for straightforward generation of synthetic emotions it serves well if it is extended with a time dependent functionality and also, with a sort of "emotional memory" which stores the most dominant emotions for a longer period of time. The latter is a feature that we plan to implement in the near future to provide more stable emotional status.

Our future work will be targeted primarily towards further refinement of the models and introducing more automated features for emotion-message act assignment. Despite the imperfect state of our current model, the incorporation of the emotional module and personality traits resulted in deeper involvement from the user side, and it also made the conversations more entertaining.

In general we can reinforce others' observations from our experience that there is a significant need for ECAs, especially on the Internet, where they could noticeably enhance the communication experience substituting for some of the power of the lost modalities.

References

Elisabeth André, Martin Klesen, Patrick Gebhard, Steve Allen, and Thomas Rist. 2000. Integrating models of personality and emotions into lifelike characters. Paiva, A. M., ed.: *Affective interactions: Towards a new generation of computer interfaces.* pp 150-165. Springer-Verlag, Berlin, Germany.

Stéphanie Buisine, Sarkis Abrilian, Christophe Rendu, and Jean-Claude Martin. 2002. Towards experimental specification and evaluation of lifelike multimodal behavior *Proceedings of the AAMAS02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'* Bologna, Italy.

Justine Cassell, and Hannes Vilhjalmsson. 1999. Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems* 2: 45-64.

Charles Darwin. 1872. *The expression of emotions in man and animals.* D. Appleton and Company, New York.

Beverly Fehr, and James Russell. 1984. Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General* 113: 464-486.

Michael A. Gilbert. 1997. Prolegomenon to a pragmatics of emotion *Proceedings of the Ontario Society for the Study of Argumentation.* Brock University, St Catherine's, Ontario, Canada.

Michael A. Gilbert. 1999. Language, words and expressive speech acts. Eemeren, F. v., Grootendorst, R., Blair, J. A. and Willard, C. A., eds.: *Proceedings of the Fourth International Conference of the International Society for the Study of Argumentation.* pp 231-234.

Katherine Isbister, and Patrick Doyle. 2002. Design and evaluation of Embodied Conversational Agents: A proposed taxonomy *Proceedings of the AAMAS02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'* Bologna, Italy.

Keith Oatley, and Jennifer M. Jenkins. 1996. *Understanding emotions*. Blackwell Publishers, Inc., Oxford, UK.

Andrew Ortony, Gerald L. Clore, and Allan Collins. 1988. *The cognitive structure of emotions*. Cambridge University Press, Cambridge, England.

Ana Paiva. 2000. *Affective interactions: Toward a new generation of computer interfaces?* Springer-Verlag Berlin and Heidelberg GmbH & Co.

Rosalind W. Picard. 1997. *Affective computing.* The MIT Press, Cambridge, Massachusetts, USA.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. Plutchik, R. and Kellerman, H., eds.: *Emotion theory, research, and experience*. vol. 1, Theories of emotion. pp 3-33. Academic Press.

Robert Plutchik. 1992. What is mood for? A critique. *Psycholoquy* 3(11).

Robert Plutchik. 2001. The nature of emotions. *American Scientist* 89(4): 344-350.

Paola Rizzo. 2000. Why should agents be emotional for entertaining users? A critical analysis. Paiva, A. M., ed.: *Affective interactions: Towards a new generation of computer interfaces.* pp 166-181. Springer-Verlag, Berlin, Germany.

Zsófia Ruttkay, Claire Dormann, and Han Noot. 2002. Evaluating ECAs - What and how? *Proceedings of the AAMAS02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'* Bologna, Italy.

Harold Schlosberg. 1954. Three dimensions of emotions. *Psychological Review* 61(2): 81-88.

Silvan S. Tomkins. 1962. *Affect, imagery, consciousness: The positive affects.* Springer, New York.

Jun Xiao, John Stasko, and Richard Catrambone. 2002. Embodied conversational agents as a UI paradigm: A framework for evaluation *Proceedings of the AAMAS02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'* Bologna, Italy.