COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

AGENT-BASED MODEL OF EMOTIONAL TRAJECTORIES IN ONLINE COMMUNITIES

MASTER THESIS

Bratislava 2012 **Bc. Maroš Gálik**

COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

DEPARTMENT OF APPLIED INFORMATICS

AGENT-BASED MODEL OF EMOTION TRAJECTORIES IN ONLINE COMMUNITIES

MASTER THESIS

THE AUSTRIAN RESEARCH INSTITUTE FOR ARTIFICIAL INTELLIGENCE (OFAI)

Field of study:	2503 Cognitive Science
Department:	Department of Applied Informatics
Supervisor:	Dipl.Ing. Dr. Stefan Rank
Licence number:	

Bratislava 2012

Bc. Maroš GÁLIK

UNIVERZITA KOMENSKÉHO, BRATISLAVA

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

KATEDRA APLIKOVANEJ INFORMATIKY

AGENTOVÝ MODEL EMOČNÝCH TRAJEKTÓRIÍ V ONLINE KOMUNITÁCH

DIPLOMOVÁ PRÁCA

MEi:CogSci Middle European Interdisciplinary Master Programme in Cognitive Science

RAKÚSKY VÝSKUMNÝ INŠTITÚT PRE UMELÚ INTELIGENCIU (OFAI)

Študijný program:	2503 Kognitívna veda
Pracovisko (katedra/ústav):	Katedra aplikovanej informatiky
Vedúci diplomovej práce:	Dipl.Ing. Dr. Stefan Rank
Evidenčné číslo:	

Bratislava 2012

Bc. Maroš GÁLIK

I declare this thesis was written on my own, with the only help provided by my supervisor and the referred-to literature and sources.

....

Acknowledgement

Foremost, I would like to express my gratitude to my super Dipl.Ing. Dr. Stefan Rank for the continuous support, his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

I would also like to thank all my friends and family for encouraging me during my studies.

Abstract

This work proposes an agent-based model of emotional trajectories in chat-like communities. We study the behavior of "collective emotions" in these communities based on data from real online interactions. Using the available data, we model and predict the emotional behavior of single users in chat-like online communities and, based on that, we also derive group behavior of a collective of users. The reference data consists of publicly available IRC-conversations collected over a time-frame of several years. The only channel that is available in order to infer internal emotional states of users is the automatically annotated content of messages authored by users in the context of the temporal sequence of texts. We propose an agent-based model in which individual emotions are reduced to two internal states: valence and arousal. In the history of an agent's communication, we observe the agent's behavior for different events – the reactions in terms of messages in different emotional states. Using these observations, we can define an agent "personality", so that each agent is a different individual or a member of a particular group. Here, an agent personality is represented by a list of n rules, where the units of a single rule correspond to expected conditions (settings) for choosing eliciting an emotional reaction in a specific situation. Using available data in order to initialize a model, we predict and investigate the "emotional trajectories" of the modeled values. The steps taken in this thesis are part of research towards modeling group behavior based on models of individual emotional response.

Keywords: agent-based simulation, cyberemotions, collective emotions, emotion modeling, emotion prediction, affective control architectures

Abstrakt

Táto práca sa zaoberá návrhom agentového modelu emočných trajektórií v online komunitách (instant messaging). Súčasne sleduje vlastnosti a správanie kolektívnych emócií, objavujúcich sa v tejto forme komunikácie. Spracovaním dostupných dát, modelujeme a predikujeme emočné správanie jednotlivých používateľov, na základe čoho môžeme odvodiť a modelovať správanie skupín používateľov ako takých. Referenčné dáta pochádzajú z verejne dostupných IRC konverzácií, ktoré boli zbierané a uchovávané počas obdobia niekoľkých rokov. Jediným možným spôsobom pristupovania k interným emocionálnym stavom používateľov počas modelovania, je emočne automaticky ohodnotený obsah správ v ich komunikácii. Toto ohodnotenie správ je robené v priamom kontexte dočasného stavu autorov pri písaní sledovaných správ. Táto práca navrhuje agentový model, v ktorom je reprezentácia individuálnych emócii zredukovaná na dva vnútorne stavy: nabudenie a valenciu. Z hľadiska histórie jednotlivých komunikácií možno sledovať správanie sa agentov v rôznych situáciách, príp. udalostiach. Najčastejšie ide o reakcie v zmysle odpovedaných správ písaných v rôznych emočných stavoch autorov. Pomocou sledovania týchto znakov sa vytvára "osobnosť" agenta. Z uvedeného vyplýva, že každý agent je iná individualita, alebo člen osobitnej skupiny správania sa. Osobnosť agenta je reprezentovaná zoznamom n-pravidiel, kde časti jednotlivých pravidiel korešpondujú s podmienkami (nastaveniami) pre výber (vyvolanie) emocionálnej reakcii u agenta vzhľadom na špecifickú situáciu. Na inicializáciu agentového modelu používame dostupné dáta. Týmto modelom následne predikujeme a skúmame emočné trajektórie modelovaných hodnôt. Táto práca je súčasťou výskumu zaoberajúceho sa modelovaním správania skupín založeného na jednotlivých modeloch individuálnych emočných odpovedí.

Kľúčové slová: agentová simulácia, virtuálne emócie, kolektívne emócie, modelovanie emócií, predikovanie emócií, emočne riadané architektúry

Table of Contents

1 Introduction

Online communities are a young phenomenon that originated in the late 80's with the advent of the internet. More recently, online communities started to grow substantially with the common use of the internet in more countries and in more homes among the general population. (2001: 500 million users; 2011: more than 2 billion users)[1]

After a few years, Internet usage became popular and useful so that “to be online” was not just a privilege of highly skilled computer users, but also the case for the people from all walks of life. For all internet users, the new communication medium also served as a way to fulfill a typical human need: to form groups - to socialize. The success of online *boards*, *forums* or *chats* was a natural result, so users had a lot of places where they could share their opinions, where to discuss their ideas, or where to just hang out for a while.

1.1 Motivation

An environment for online communication substantially differs from a real-life one. In an online environment, users can often communicate only with text, but nevertheless try to communicate a whole range of emotions. Online users are simply not able to hear a tone of voice of a sentence of another user and other non-verbal cues that enable recognition of irony, sadness or happiness. Users are not able to see another user's facial expressions or gestures during a dialog. On the other hand, internet users supplanted this disadvantage by the usage of emoticons (internet smileys: *:-) :-(*), which have an impact on message interpretation and are useful in strengthening the intensity of a verbal message (Derks, et al. 2008).

In the recent past, this new phenomenon became a point of interest for research groups (Schweitzer and Garcia, 2010; Chmiel and Holyst, 2010; CyberEmotions, 2012). In February 2009 a consortium called CyberEmotions (CyberEmotions, 2012) began its work in a research domain that studies collective emotions in an online environment. It focuses on: *“observable and analyzable phenomena related to any means of communication*

1 Internet Telecommunication Union. 2012, February 28th. Global numbers of Internet users, total and per 100 inhabitants, 2001-2011. Retrieved from http://www.itu.int/ITU-D/ict/statistics/

provided by the Internet - such as text, sound, visual, or any combination of these - that are related to emotional processes in individuals or groups.“ (CyberEmotions, 2012)

Due to the growing number of people that gather online, many commercial companies started to be curious about “what's going on in internet”. Considering the reaction of users of online communities to a special event such as the release of a new product from a company expected to set new trends, is a good example for how a prediction of such an emotional reaction of people to a new event could be important and useful. One such example happened on June 25th 2009 when a very famous pop singer, Michael Jackson, died. Internet users started to look for news about his death. The biggest web pages reported a massive increase of page views in one hour: *“CNN reported a fivefold rise in traffic and visitors in just over an hour, receiving 20 million page views in the hour the story broke”*, *“Between approximately 2:40 p.m. PDT and 3:15 p.m. that day, some Google News users experienced difficulty accessing search results for queries related to Michael Jackson,”* (Rawlinson, 2009). A prediction of such events could prevent a potential traffic overload by a previous preparation for this kind of situations.

In the example above we can see that an activity of a group of online users can foretell or predict much more reactions from other users. This phenomenon occurs especially when users' activities are emotional. Observing emotional communication and its effects on a group could lead to predictions of such effects and behaviors of groups in the future.

One way to simulate a group's behavior is to simulate the behavior of every individual in this group, while deriving group behavior from an aggregate of these individuals. Moreover, simulation of individuals could be useful as a support for 3rd parties (agents, robots, companies) for knowledge of a possible impact on individual's reactions of the 3rd party's actions.

In this work we want to create a model which will simulate emotional behavior of an online group and will be able to do simple-and-quick predictions of the group's individuals.

1.2 Datasets

The data that we rely on as a reference for our work are the logged chat communication on the support channels for the operating system Ubuntu. Ubuntu is a Linux computer operating system which is distributed as free and open source, and used daily by more than 20 millions of users[2]. The protocol used on these chat channels is Internet Relay Chat[3] (IRC) . IRC is one of the oldest protocols used on the internet for communication channels that connect groups of users allowing them to talk to each other synchronously. IRC provides two forms of chatting: "public" and "private". The public form is represented by named *channels* which are mostly moderated rooms where a group of people can talk with each other. The second form is private messaging between two users.

The reference data we use consists only of the public channels' communications, so we do not have access to any private communication that happened simultaneously. This is an unavoidable factor for any such data since users could be influenced by their private communications as well as by other elements (e.g. users' communication offline, external events, …).

The subset of the Ubuntu IRC dataset we use as reference for the model described here consists of a full history of all the communication on 3 IRC channels: from 2008-01-16 to 2007-07-18 for the IRC channel #ubuntu-irc, from 2006-08-23 to 2010-07-16 for #ubuntu-laptop and from 2008-07-10 to 2010-07-18 for #ubuntu-website. The purpose of most of the logged IRC channels is to provide a forum for support or troubleshooting for Ubuntu users as well as for developers.

The logged communication is explicitly published in the public domain. Other similar datasets as well as any real-time logging of online communication, however, have to respect the privacy of users, i.e. to regard all communication of IRC users as private by default. Therefore, our modeling and simulation effort does not have direct access to the whole dialogues word-by-word. The data is pre-processed by an anonymizing tool that replaces human-readable user ids with numerical ids and replaces the actual utterances with annotations of the utterances based on available classifiers and language processing tools described below.

What features are available in the anonymized data?

- A **sender** and potentially a **recipient** of a message. For every single message, there

2 Ubuntu. 2012, 2nd April. Retrieved from Ubuntu homepage: http://www.ubuntu.com.

3 http://www.irc.org

is one and only one sender of the message. As IRC does not allow to send messages to channels anonymously, the sender of each message is always known. On the other hand, the recipient(s) are not always directly known, as users do not need to explicitly mention the recipient of a message. By the default, a message has no explicit recipient. However, one of the annotation tools provided by the CYBEREMOTION project and used to annotate that dataset, uses a heuristic based on common practice in IRC communication (mainly prefixing of utterances with another uses nickname) to derive a single recipient from the original data if possible.

– **Current user's emotions:** for every user's utterance there is a related emotional evaluation of every single utterance. Briefly, an emotional evaluation estimates a positive or negative polarity of an utterance as well as a value that corresponds to the level of affective involvement of the author. These values are based on two different classifiers: *SentiStrength* (Thelwall, 2010; Thelwall, 2012) and *ANEW* (Bradley and Lang, 1999). (See Chapter:2.2.6).

– **Further annotations:**

  – **Event type:** defines a type of an event in an IRC channel, whether it is a user's normal message to other users or some special action (e.g.: nickname change, channel join, channel left, channel's topic change)

  – **Message length:** length in words and characters

  – **Message dialog act class**: classifies a message into one of the categories: greeting, acceptance/rejection, emotion, emphasis, question/answer and statement

Based on this set of features, we model the affective communication between users and their emotional states in terms of trigger – action content. For every human's emotional action, no matter whether verbal or non-verbal, there must be some object, person or situation which elicited it (Cowie, 2007) . However, in the online environment we are not able to determine all potential real-world triggers of an action. What we can collect and analyze are all users' online actions and users' emotional states (in a way we define them).

Therefore for every agent's action we try to find its trigger action according to senders and recipients of previous messages, it's time position in a communication and agent's emotional state. As we are not able to always find a trigger of an action, some triggers remain unknown. This approach is the main idea and motivation of the work.

1.3 Terminology

In this work, we use terminology used in psychology such as *emotion*, *arousal, appraisal, mood, feeling, emotional state*. We also describe several emotional theories by different authors that do use the same terms. Different perspectives (e.g. scientific, common language) use different definitions of the terms. This is a natural consequence of an overlap between everyday understanding and scientific theories of emotions in psychological and neurological research (Marsella, Gratch and Petta, 2005). As we base our model on psychological emotional theories, we use psychological definitions of what emotions and its elements are. See Chapter 2.2 for the specific definitions of these terms as far as they are relevant for this work.

1.4 Hypothesis

Based on annotated data of a history of discussions in an online community, a model of the emotional behavior of (groups of) individual users in these communities can be constructed. Predictions of future behavior of individuals based on this model are consistent with the reference data for the same period as well as with mathematical models of similar communities based on aggregate measures after a long-enough period of observations. The model's predictions of the future behavior are superior to predictions from a stochastic model which is based solely on observed event probabilities.

1.5 Goals

The model developed in this thesis has to balance the need for a suitable degree of complexity with the need for a simple and fast simulation. The latter is a consequence of one of the application areas for such a model: as online decision-support for interactive conversational systems.

1.6 Interdisciplinarity

Modeling emotions of users in online communities is a recent topic coming from different fields of study. As the topic includes social (human) research in emotions as well as computer based modeling, there is not a single scientific discipline covering all needed methods. This phenomenon requires multiple scientific approaches to be involved to fully understand the topic and to be able to achieve the goals. The interdisciplinary approach for this work involves research in psychology, which is concerned with the internal life of a single user (agent) including human emotions and personalities. For the perspective of single user behavior in the group and the behavior of a group as a whole, sociology takes on an important role. To establish a theoretical perspective and to ground the psychological and sociological elements of this model, modeling tools from physics are used. For the practical part of this work: it's implementation and modeling itself, computer science is used.

1.7 Structure of the Thesis

The thesis is structured into 5 main chapters including this chapter, introduction, where we describe our motivations, hypothesis and goals that we want to achieve. After introducing the issues of an *Agent-based Model of Emotion Trajectories in Online Communities* in chapter 1, the second chapter presents *Related Work and Theory.* We firstly define and describe two main approaches used in the work, *Agent-based modeling* and *Affective Computing.* In the end of this chapter, we survey previous work that has been done in the topic. After the two theoretical chapters, we suggest our implementation of previously mentioned ideas in Chapter 3, which also includes the formal definition of the proposed model. The fourth chapter, *Results*, describes an evaluation of the model and discusses the results gained from the model simulations. The fifth, and last, Chapter summarizes the work, presents its conclusion and suggests future ideas about further work or its improvements.

2 Related Work and Theory

The modeling of individual emotional behavior has been an active research topic for a substantial amount of time (Grach and Marsella, 2001; Marsella, Gratch and Petta, 2005; Moffat, Frijda and Phaf, 1993). This chapter overviews affective computational models and related theory. The chapter is divided into three main parts: Agent-based modeling and simulation, Emotions, and Previous Work. The first part is dedicated to theoretical computer science models of agents, agent-modeling and multi-agent systems. The second part introduces a psychological phenomenon called Emotions, emotion modeling from computational perspective and important appraisal theories which are essentials for the work. The last part describes the CYBEREMOTIONS project (Cyberemotions, 2012), in the context of which this work has been performed, and mainly Schweitzer's work (Schweitzer and Garcia, 2010) *An agent-based model of collective emotions in online communities* which focuses on the same topic, but from a different perspective.

2.1 Agent-based modeling and simulation

Before any introduction of what agent-based modeling is, we should define a concept of *agent*. *Agent* is a term used in computer science, especially artificial intelligence, for describing an artificial autonomous process placed in an environment, which repeatably perceives its inputs and acts based on them in order to achieve its goals.

(Nwana, 1996) says that agents may be classified along (at least) three ideal and primary attributes which agents should exhibit: *autonomy, learning* and *cooperation.*

Autonomy of an agent is its ability to operate in an environment independently without any human intervention, even if the environment is not fully observable. Agents act in a manner to follow their internal states and goals. *Cooperation* stands for a group of social abilities by which agents can communicate with other individuals: agents or possibly humans. The agent's cooperation ability is a main reason for having multiple agents in an environment instead of having just one. *Learning* is a avery important capability by which an agent can react and interact **differently** over time in their dynamic and indeterministic environments.

Even though there are big efforts in agent-related research, there is not one generally accepted definition of an *agent*, however, the most quoted one is from Wooldridge & Jennings:

„An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives.“ (Wooldridge, 1995)

An *environment* is a set of all objects that an agent interacts with, comprising everything outside an agent. An agent perceives input from an environment by its sensors and performs actions by its actuators.

Now, we can introduce a *single-agent system*. In a single-agent system an agent models itself, an environment and its interactions. If there is another agent it is considered as a part of the environment.

One could assume that a system built on a single-agent would be necessarily simpler than a system built on multiple agents. This, however, is not necessarily true, as multiple agents allow to separate a complex problem into smaller parts, agents. Hence agents can be modeled as simple systems with specialized interactions among them.

2.1.1 Multi-agent System

A multi-agent system (MAS) is a system of multiple agents based in an environment interacting among themselves trying to achieve goals. In such systems every agent has it's own sources, internal states and limits by which it can freely act, but on the other hand it is also limited by them and its capabilities. In an MAS, every agent has its own place and goal which it tries to achieve. In such an "agents society", agents coordinate and communicate with each other in order to arrive in the end at a collective solution of a problem for which the MAS was designed.

Paraphrasing (Wooldridge, 2009), agents can represent behavior of individual people or individual agents can represent organizations and groups. Thus an MAS model consisting of such agents can be used to simulate the behavior of human societies. Additionally, (Gilbert and Conte, 1995) suggests that multi-agent simulation of social processes can have the following benefits:

- *"computer simulation allows the observation of properties of a model that may in principle be analytically derivable but have not yet been established;*
- *possible alternatives to a phenomenon observed in nature may be found - properties that are difficult/awkward to observe in nature may be studied at leisure in isolation, recorded, and then 'replayed' if necessary;*
- *'sociality' can be modeled explicitly - agents can be built that have representations of other agents, and the properties and implications of these representations can be investigated."* (Wooldridge, 2002, p. 260))

Another argument for using an MAS model to simulate social behavior is from (Moss and Davidsson, 2001, p. 1):

„[For many systems,] behavior cannot be predicted by statistical or qualitative analysis. ... Analyzing and designing ... such systems requires a different approach to software engineering and mechanism design."

Based on the arguments mentioned above, we decided to use a Multi-Agent-Based System as the main modeling approach for our work.

2.2 Emotions

Emotions are part of our everyday life, everyone has them, lives with them, but not a lot of us really understand them. In this subchapter we try to define and describe emotions from a psychological perspective as well as make an overview over important emotional models and theories that are important for this work.

Emotion is a wide ranging term used in everyday language, without a clear universally accepted definition. There are two types of definitions of emotions: *Experiential* and *Conceptual.* The *Experiential* type is couched in terms that can be observed by ordinary means:

"a state where the person's whole system is caught up in the way they react to a particular person or situation – which may be in reality or in their mind. Specifically:

- *it involves distinctive positive or negative feelings about the people or situations involved*
- *it involves impulses to act or express yourself in particular ways and avoid others*
- *it involves distinctive changes in your body, for instance in your heart rate or tendency to sweat*
- *it doesn't usually last very long - it comes on quite quickly, and dies down reasonably soon (unless there is something very unusual happening)"*

(Cowie, 2007, p. 19)

The *conceptual* interpretations are couched in terms that suppose *to lie behind* the observable patterns, and only way how to observe them is by specialized techniques (technological or inferential).A widely supported conceptual definition from Scherer is:

„episodes of massive, synchronized recruitment of mental and somatic resources allowing to adapt or cope with a stimulus event subjectively appraised as being highly pertinent to the needs, goals and values of the individuals"

(Scherer, et al. 2004)

The experiential definition is relevant for clarifying the way how the term of emotion is used in everyday life, on the other hand the conceptual one is usually tied to some theory. As we base our computational model on already proposed emotional theories, in this work, we work with a conceptual definition of emotions.

Emotion is a complex psycho-physiological experience, which has a big impact on human behavior. It influences our conscious as well as unconscious processes, such as: behavior, perception or memory. One of the aspect of the emotion influence of declarative memory is that events with any positive or negative emotional value are often more remembered than events that are emotionally neutral. On the other hand, the majority of studies that compared influence of declarative memory by valence and arousal of an event left unanswered the question about exact relation of the valence dimension and arousal dimension to the memory enhancement effect (Kensiger, 2003). Thus, the model that we propose should distinguish between a neutral emotional event and a positive or negative one. The differentiation will lead the model to select a different mechanism of storing (remembering) to its memory.

Conceptual accounts of human emotional phenomena are done by two major methods in psychology. The first is a discrete approach while the second is a dimensional one. The discrete approach is based on the direct language analysis where the emotional words and emotional expressions are related to separate emotional states. This approach relies on a semantic categorization to semantic groups which represent “basic emotions” (Ekman, 1999). Shortly, “basic emotions” refers to a theory which claims that types of human emotions can be separated into a finite number of emotional types. During that psychological experiments on emotions, participants identified the expressions of emotion in photographs of people from different cultures. In other cases, the participants have to indicate an emotionally relevant evaluation which they remember feeling during the experiments (Scherer, 2005).

The second approach is dimensional. The first recent dimensional theory was proposed by Wilhelm Wundt in (Wundt, 1880) where he provided a method of describing subjective feelings (states) by a 3-dimensional model. The three dimensions are: “pleasure vs. displeasure” (valence), “rest vs. activation” (arousal), “relaxation vs. strained attention” (tension). Wundt with his theory influenced the later research in emotion measurement (Schlosberg, 1954). Because of the difficulty of identifying a tension, control or potency, some theories were reduced to a two-dimensional approach omitting the third dimension (Scherer, 2005). On the other hand, some emotional models still use the three dimensional approach (Mehrabian, 1996). Using a 2-dimensional approach reduces the possible types of situations (emotional events) by which an a computational model stores the information about the situation. To say it simpler, – using the 2-dimensional approach simplifies the possible model by reducing its complexity (and increasing its robustness) which leads to a simpler detection of emotions and in our case increases the performance of the model itself. Adding a 3rd dimension into the 2-dimensional valence&arousal could extend the possible size of a computational emotional model very rapidly.

The reason that we do not use a discrete approach is that we do not have an access to an equivalent of participants' self-reports in IRC, however we do have an access to their textual expressions from which we can measure their arousal & valence. Therefore a dimensional approach is used for emotional analysis of IRC chats.

Another argument for using a dimensional approach is stated in (Mehrabian, 1978): *"From a measurement standpoint, emotional traits (or temperament) can be inferred by measuring and averaging an individual's emotional states across a wide and representative sample of everyday situations."* Here *emotional traits* corresponds to a term that we defined as a *personality.* The values of a discrete approach with basic emotions emotions (anger, sadness, happiness, relaxation) can not be directly averaged, therefore we need an approach with ordinal or continuous values of emotional states. Additionally, Mehrabian argument also supports our approach of creating a *personality table* by measured values that are obtained by observing an agent's emotional states and reactions in a representative samples of everyday situations.

2.2.1 Suitable dimensional approaches

Emotional experience can be described by two factors: valence and arousal. V*alence* represents how negative or positive the emotional experience was for a person, while *arousal* describe how calming or exciting it was. On the following Figure 1, you can see a two-dimensional representation of emotion in a semantic space. On the X-axis, the value of arousal is presented a on the Y-axis the value of valence. The values of valence and arousal can define a position in the space semantic of emotions. In the figure you can see, that the space is divided into the smaller parts, which represent an emotion.

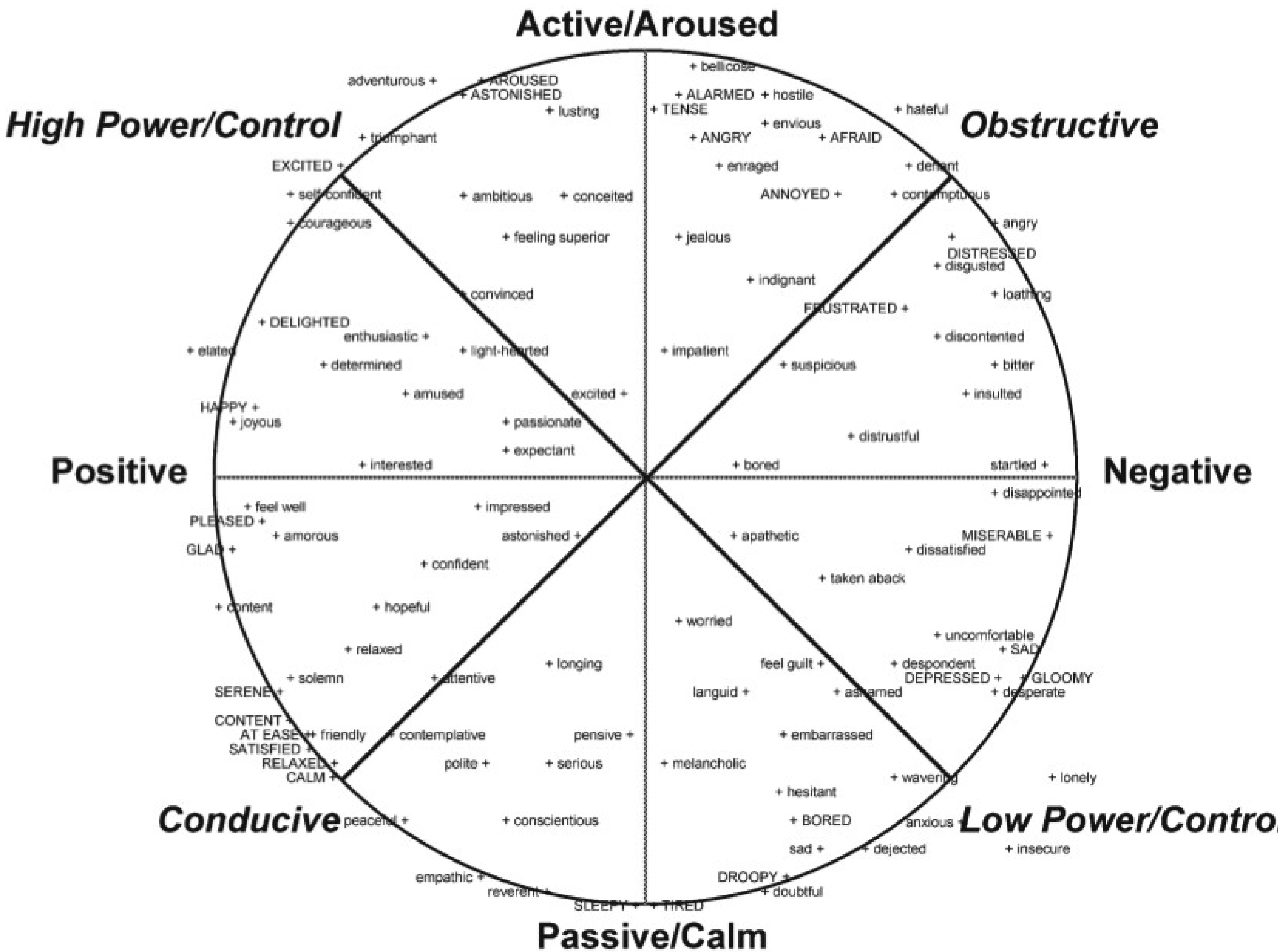

Figure 1. Alternative dimensional structures of the semantic space for emotions (Scherer, 2005)

Arousal

In psychology, *arousal* refers to an awakening state of a person to any stimulus (Cowie, 2007). In our work, we use this term to describe a user's effort in writing (expressing) his reactions (feelings) in an online environment.

Valence

In psychological emotional theories, *valence* generally represents a positivity or a negativity of a character of an emotion (Cowie, 2007). In our work, we use valence as a positive-negative polarity measurement of a written (expressed) by a user in an online environment.

For our purposes we decided to use a two dimensional model of emotions (e.g., valence & arousal). As mentioned in Chapter 2.2, the number of dimensions in the model of emotion could have negative impact on the size of computational model. Additionally, when we try to match two emotions in a dimensional model, the number of dimensions is essential in a "matching" complexity as well as the number of dimensions can influence the correctness of such matching. We mentioned that because the matching of two emotions in time is crucial for our purposes as we assume that humans behave similar under same conditions. Such assumptions are based on the following theories.

2.2.2 Appraisal Theories

General description of appraisal theories.

Appraisal theory is based on the idea that emotions are elicited from our evaluations (appraisals) of events or situations that cause specific reactions in different people. This was first introduced by Magda Arnold in (Arnold, 1954). Since then, many appraisal theories have been proposed (Roseman, 2001).

Appraisal theories are theories which are most used (and potentially most suitable) among psychological perspectives on emotion as main source for designing computational systems. The main reason is that appraisal theories explain the connections between and more clearly define the processes involved in cognition and emotion (Marsella, Gratch and Petta, 2005). These theories differ in several aspects of how an emotion is elicited according to a stimulus, but one common thing that they all share is that they are all based on the idea of how individual differences in affective responses can be elicited by the same stimulus, what explains why two different people can significantly differ in their emotional response to the same event (Mascarenhas, 2001). This is a very important point, as we model collective behavior of different single individuals instead of modeling the behavior of single individuals with identical personalities like in (Schweitzer, 2007) .

Coping

Coping is an essential part of human emotion. It describes the process of dealing with specific situations under specific settings that were appraised as relevant to the individual. In appraisal theory, it is assumed that a repeated emotional situation with the same parameters can elicit the same emotional reaction as it did before (if all the necessary settings are the same). Different types of coping can be distinguished:

"*Embedded in the Ways of Coping scale is a distinction between two general types of coping. The first, termed problem focused coping, is aimed at problem solving or doing something to alter the source of the stress. The second, termed emotion focused coping The second, termed emotion focused coping, is aimed at reducing or managing the emotional distress that is associated with (or cued by) the situation. Although most stressors elicit both types of coping, problem-focused coping tends to predominate when people feel that something constructive can be done, whereas emotion-focused coping tends to predominate when people feel that the stressor is something that must be endured (Folkman and Lazarus, 1980)*." (Carver, 1989, p. 267)

As the main idea of our work is to predict an individual's emotional reaction according to the history of individual's behavior, modeling of Coping behavior (Marsella, 2003) is important part of this work.

Personality

A nice definition of the personality that serves well for modelers and generalizes the concept is:

"*Given an agent with certain functions and capabilities, in a world with certain functionally relevant opportunities and constraints, the agent's mental reactions (behavior,* *thought and feeling) will be only partially constrained by the situation it finds itself in. The freedom it has in which to act forces any action to reveal choice or bias in the agent, that may or may not be shown in other similar situations. Personality is the name we give to those reaction tendencies that are consistent over situations and time."* (Moffat, 1997)

Another nice definition is from (Revelle and Scherer, 2010), where a personality is defined as the coherent patterning of affect, behavior, cognition, and desires (goals) over time and space. A helpful analogy for understanding the differences between emotion and personality is: „*is to consider that personality is to emotion as climate is to weather. That is, what one expects is personality, what one observes at any particular moment is emotion.*" (Revelle and Scherer, 2010)

Emotional Life is led by (whole) persons who experience and cope with emotions. One type of a current research in personality can be organized by level of generality between people (Revelle, 1995). It ranges from generalizing to all people to focusing on single individuals how:

- **all** people are the same, i.e. species-typical behavior;
- **some** people are the same, individual differences and similarities (e.g., traits);
- **no** people are the same, but show unique patterns of behavior. The emotional life, as proposed above, perfectly suits to the behavior of users in online communities and describes the behavior of possible computational agents that model it. Specifically:

Hence, this concept of personality is consistent with our use of personalities for a computational model of users emotional behavior in online communities.

Structural vs Process Appraisal Theories

A common classification of appraisal theories is based on a structural versus a process-based description (Roseman and Smith, 2001). As (Reisenzein, 2001) notes, it is important to keep a clear distinction between the process of making appraisal and the outcome of this appraisal process. Appraisal theories, called structural theories, claim that people appraise objects by dimensional evaluation; the theories suggest that it is possible to examine an individual's appraisal of a situation and then predict the emotional experiences of that individual based upon his or her views of situation. Structural dimensional theories of appraisal assume, that humans somehow assign appraisal values (by dimensions) to an object. However this assigning of values works, it is not defined by those theories and it is supposed to work as a “black-box”. Because the function of assigning values of appraisals is not defined in the structural theory, any way of determining these values is feasible.

The structural theories of appraisal describe the structural relations between appraisals (dimensions) and elicited emotions. The task of assigning appraisal function values to input objects is the task of appraisal process theories. On the other hand, structural theories can do that without any process assumptions. Structural theories have been criticized for failing to capture the dynamic nature of emotion. For the deeper analysis of appraisal complexity, psychologists thought about further complementing the structural model by understanding the appraisal as a process. Process theories rely on the idea that it is important to specify the cognitive principles and operations that are behind appraisals.

Several process theories have been proposed (Smith and Kirby, 2000; Smith and Lazarus, 2001; Scherer, 2001). As an example of a process appraisal theory we take Scherer's (Scherer, 2001). The reason that we do not use Scherer's appraisal theory directly in the model is, that it is a *Process Model* theory which is made up of three levels of appraisal process, with sequential evaluation checks at each level of processing that create a specifically ordered processing construct (Scherer, 2001). In our case, we cannot use this theory as basis for implementing the determination of an emotional state directly as the information in the interaction data is not complex enough to identify all necessary checks for every interaction. Rather we use an n-dimensional table in which we do access information through the “dimensions”, so there is no influence of the “dimensions” order on the result of emotional state of an agent. According to this, we use a *Structural Model* appraisal theory. Furthermore, other aspects of the reason to use a Structural Model appraisal theory, are that we do not have any access to agent's physiological changes, nor to a sequence of agent's physiological changes right after a new event occurs, what is one of parts of Scherer's Appraisal theory. Here, we can only observe the agent's expression to a new stimulus. For better understanding the reason of using a structural theory and dimensions, one can look at (Ellsworth and Scherer, 2003) where it is proposed that appraisal theories assume that the type of emotion elicited by an event can be reliably predicted if it is known what kind of conditions (settings) appraised the event. Hence, we can represent the result of the appraisal process as a profile of evaluations on the basic appraisal dimensions. Based on this several appraisal theorists proposed theoretical predictions about profiles (appraisal criteria) of some basic emotions. The following *Table 1* illustrates this approach in form of a generic prediction table. The first column represents dimensions (appraisal criteria) and others holds the value needed for eliciting a specific emotion.

Appraisal Criteria	Joy/Happiness	Anger/Rage	Fear/Panic	Sadness
Novelty	high	high	high	low
Intrinsic pleasantness	high	open	low	open
Goal significance				
Outcome probability/certainty	high	very high	high	very high
Conduciveness/consistency	conducive	obstructive	obstructive	obstructive
Urgency	low	high	very high	low
Coping Potential				
Agency/responsibility	self/other	other	other/nature	open
Control	high	high	open	very low
Power	high	high	very low	very low
Adjustment	high	high	low	medium
Compatibility with standards/ value relevance/legitimacy	high	low	open	open

Table 1. Example of Postulated Appraisal Profiles for Different Emotions (Ellsworth and Scherer, 2003)

2.2.2.1 Frijda's emotional theory

According to (Moffat, 1997) the best computer-implementable theory is Frijda's emotional theory (Frijda, 1986). Frijda's theory falls under the group of appraisal theories. It identifies emotions as changes in activation of behavioral dispositions, caused by relevant stimulus events. (Moffat, 1997) In this theory, *concerns* occupy the central position in an organism which represent its needs or preferences. They determine which situations are processed and which are filtered out, so an organism does not deal with them anymore.

The theory proposes that emotions arise as a result of two-stage appraisal of the stimulus: primary and secondary appraisal. Primary appraisal is a continuous relevance detection procedure to see if a stimulus is matching against one or more concerns. Primary appraisal matches events identifying positive or negative feeling from the previous emotional experience (Moffat, 1997). Secondary appraisal continues on an event relevance detection from a contextual point of view. The appraisals given by the theory include identifying features of the surrounding situations according to the organism's goal and preferences. e.g. whether an event is *conducive* or *unfavorable;* its *fairness (honesty* or *morality)*; who is a responsible author of an event; suddenness or (un)expectedness of the event, etc. (Moffat, 1997)

"It is significant for the latter appraisal category what responses are available to the organism to cope with the event, how costly they are, and how feasible or likely to succeed." (Moffat, 1997)

The primary appraisal refers to the potential of the input stimulus perception to be noticed by a mental processing of an organism. This is called *control precedence.* When an organism's internal processing attends to an output stimulus, its internal states are changed and then it performs an action as a result of this change. The change of *action readiness* refers to the emotional responses which are provoked by the stimulus situation (Moffat, 1997). *Action tendencies* can be seen as first impulses that organisms feel immediately after an emotional event, e.g. to approach, hurt, flee, attack, defend. This *action tendency* represents how an organism would approach or avoid a situation without any other further processing. The secondary appraisal processes the contextual features of an input stimulus and regulates an organism's determining of possible action that he wants to execute.

For better understanding, imagine a situation where you got an unexpected strong hit from someone who you respect or love. Your first action tendency is a possible return of the hit, cursing or any other kind of revenge. However, the procedure of the secondary appraisal results in not striking back, because it could be something that you can regret in the future.

From the point of individual emotional behavior, a good work laying out an overview of computational models of emotion is (Marsella, Gratch and Petta, 2005) which contains an overview, history and comparison of them.

Appraisal Emotional Models

Two models were proposed based directly on Frijda's theory: ACRES (Frijda and Swagerman, 1987; Swagerman 1987) and WILL (Moffat, Frijda, 1995; Moffat, 1997). The first model was ACRES which main propose was was the need for an organism to continually check all its inferences from what it perceives, and from what it believes, and from what it predicts, and from what it intends (Moffat, 1997).

WILL

Will is a successor to ACRES. The motivation to create Will was to have as simple model as possible which uses existing AI technologies.

The model is an autonomous agent with a certain problem-solving mechanism and inference capabilities to support its autonomy. As the goal to create such an agent was to have a model that would be easily implementable in a computer program (an automatic system), the emotion processing which the model represents is disassembled into smaller parts which enables them to be implemented by the current AI methods. However the model is transformed, it must be consistent with Frijda's emotional theory. This approach leads to a final model which covers a lot of emotional issues in AI and is much like a standard symbolic system[4](Moffat, 1997).

(Moffat, 1997) broke down the theory of emotion into more elementary cognitive elements of emotions. He stated that only few elementary types of cognition are necessary.

In the emotional process, one need to firstly perceive any stimulus (emotional stimulus) to be able to process an emotional event. An independent module called ***Perceiver***. The complexity of *Perceiver* can be very complex system as a stereoscopic vision or very simple as a keyboard or any simple text input is.

As the *Perceiver* is responsible for the model inputs (perceptions), a second independent module called ***Executor*** which executes actions which are described or shown in the output interface of the module. (Moffat, 1997) argues that this simplistic split off perception from execution is done in the way that used to be dominant in AI in the time the model was proposed (1997) and not a better solution was proposed yet.

Concerns module is the first and primary appraisal of the stimulus event. This module matches an input event positively or negatively (or it's affect polarity) against a previous history of emotional experience.

According to Frijda's emotional theory, the next part of the stimulus event processing is

---

4 A computer program that performs computations with constants and variables according to the defined rules. Symbols (constants, variables) and their meanings are not grounded and the system has *de facto* no idea what those symbols mean. Symbolic system just manipulates with those symbols without any grounding to the meanings

context evaluation (secondary appraisal). A module called ***Predictor*** evaluates specific properties of events. As the concerns module takes care about emotional value of a stimulus event, and reacts according to an agent's needs and preferences, *Predictor* process the context (features) such as: author and receiver of an event, timing, clarity vs. ambiguity, etc.

The theory claims that the next step that precedes the emergence of emotion is an evaluation of seriousness of a stimulus event. The seriousness of a threatening situation differs according to its reality and its level of danger for you. For easier understanding, imagine that you are in a wood and you see a big mighty snake. If you are not a snake buster, you probably would be very scared and try to run away as fast as possible. On the other hand, if you are still not a snake buster, but the snake is closed in a cage, you would be not so scared and finally you would evaluate the situation in a very different way even thought that the setting that you are in a wood with a snake is still the same. An another module which solves such part of an emotional development is called ***Planner***. Problem solving is a part of AI which can be implemented into an autonomous agent.

In every day life, we often meet with situations or events that somehow evoke emotions in us, but we do not understand why such emotions were elicited. (Moffat, 1997) uses this example: *"Even when people react angrily and hit out at strangers, it is not clear what they are trying to achieve."* So, if there is no conceivable goal, which is the case with much emotion, then the modules introduced so far are not enough to account for emotion. The remainder of the emotion process, that cannot be explained by the modules, will be confined to its own module, the ***Emotor,*** which is explained later. (Moffat, 1997)

By other words, the main part of emotion is modeled by the modules described above, including affect, primary appraisal and control precedence. The other part of emotion, secondary appraisal and action tendency is modeled by the *Emotor*. Imagine a situation of the anger emotion, by which one has an intention or tendency to harm someone or destroy something because of immoral action of his. This shows us that emotions are somehow arbitrary. The part of *Emotor* can be set (programmed) to elicit some action tendencies in case of some appraisals.

The module through which all the other modules communicate is called ***Memory.*** The memory is responsible to store all the information about previous and new appraisals, new experiences, agent's reactions and simple the all input-output information that flows cross the model.

For better understanding (imagining) of how the model looks like, you can see a following Figure 2.

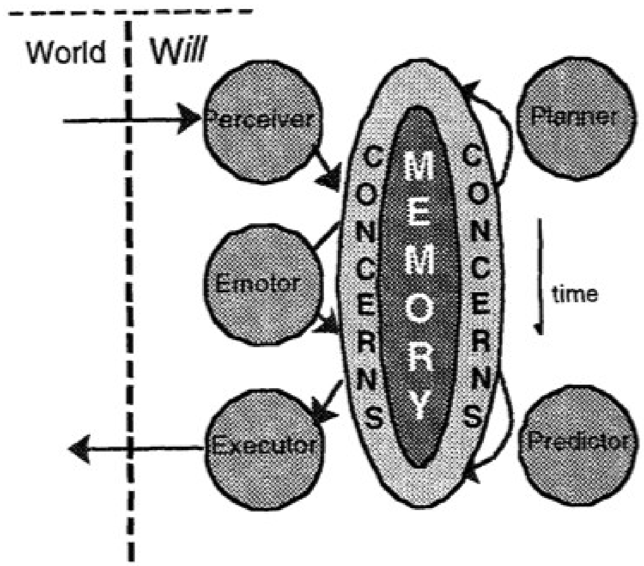

Figure 2. Will - Full Architecture, taken from (Moffat, 1997)

We presented the model of Will as a good start point (basis) for creating and a developing an agent model suitable for our purposes. The model is based on Frijda's appraisal theory which is applicable for the problem targeted in this work. The architecture of the model also serves us as a good prototype for choosing a suitable implementation of the model.

2.2.3 Combination of Emotions

As we already introduce, we model an affective communication in the sense of trigger-action. After any agent reaction in online environment we need to find a trigger (reason) which elicited an agent to act. Relying on the data we use, we have an access to a non-complete heuristic annotation of messages recipients. Therefore we are able to find a single trigger message (event) of an agent reaction. However, humans emotional reactions are not always related to only one emotional event, or not only one emotion is related to an event. Such a behavior also occurs for online users. The current model does not cover situations when multiple triggers precedes an agent reaction. A good example of such IRC communication is the following :

12:20:03 <Jane> John, I think this is not the best place for you

12:21:32 <Mary> John: I think so too. Don't be stupid...

12:23:47 <John> You all would like to get me out of here, but I don't care...

Text 1.Example of IRC communication

In the example above (Text 1) you can see that it is simply not possible to distinguish what elicited John's reaction between the two single-event triggers. Here a combination of previous emotional events comes into the role as the John's reaction was a result of combination of both previous messages (Jane's and Mary's). To improve a model to not miss the ability of combining emotions for a trigger detection, we need to compute the overall emotional intensity of a number of emotions. Hence we decide to implement and propose a feature for the model, which does not handle only single-event triggers, but also multiple-events triggers.

Imagine a situation when your mother-in-law kills herself in a car crash in your new car. One can feel sadness because of his mother-in-law died and maybe anger because she drove without his permission. Here we have two emotions that could be elicited by this bizarre situation. Let say that the emotion of sadness has an emotional value 0.7 and the anger emotion has 0.5. The interval of this emotional value is from interval <-1, 1>. How such a combination of two distinct emotions can be combined in a final intensity? (Reilly, 2006) postulates these basic rules for a combinations of more emotional intensities:

The combination is **non strictly addictive** i.e. 0.5 + 0.7 does NOT equal 1.2 (we can not use addition). Also **multiple emotions should each play some role:** 0.5 + 0.5 does NOT equal 1 (we can not use maximum function). In the end, a **result should be at least as intense as most powerful emotions**, so 0.5 + 0.7 does NOT equal 0.6 (we can not use an averaging function).

(Reilly, 2006) presents two proposals that hold the three conditions above: *logarithmic combination* and *sigmoid function* (Picard, 1997) Where the both stands approaches hold the above conditions for combing emotions. For our purposes we use the **Logarithmic combination** of emotions,

$$f(em) = 0.1 \times \log_2 (\sum_{em} 2^{10 \times intensity(em_i)})$$

Figure 3. Combination of emotions - logarithmic combination

where the graph of the functions f_v *and* f_a are linear near 0, less so approaching 1.0,

f_v stands for the valence part of internal state and f_a stands for it's arousal part,

em is a list of emotions,

em_i is an individual emotion from the list *em,*

intensity is a numerical value of a dimension (valence or arousal) for the emotion.

To be able to group (combine) multiple triggers for into a single emotional event, we use one the logarithmic combination proposed above. The grouping of triggers into a single unit leads into the having a simple computational model based on a trigger-reaction mechanism.

2.2.4 Collective Emotions

One of the goals of this work is to be able to capture collective behavior on the internet. When a number of users communicate their feelings together in the environment at the same time a phenomenon called *collective emotions* appears. To measure an intensity of collective emotions we use the concept of *collective emotional state.* Collective emotional state emerges if a sufficient number of agents expresses their individual valences at a specific time. Dealing with the emergence of collective emotions in online communities is the goal of CYBEREMOTION project (CyberEmotions, 2012) which focuses on understanding *„the role of collective emotions in creating, forming and breaking-up Information and Communincation technologies mediated communities as a spontaneous emergent behaviour occurring in complex techno-social networks.“ (CyberEmotions, 2012)* Good source reference for understanding Collective Emotions and their emergence on collective behavior can be found in publications of CYBEREMOTION consortium e.g. (Chmiel, et al. 2011a; Chmiel, et al. 2011b).

Collective Emotional State

We define the collective emotional state as a sum of all absolute values of emotions in a specific environment for a specific time window. We take into account also the number of people expressing their emotions and the number of messages (expressions) by which these emotions were shared among the environment. Thus, we can define a collective emotional state by 4 values: total valence, total arousal, number of people, number of messages. We decided for absolute values of valence & arousal because of different kinds of how people reaction on some triggers. As some reactions can elicit strongly positive emotions, for other people it could lead into intensively negative ones. By simple summing all those emotional values, the final value could be a value near 0, because positive and negative values can simple neutralize themselves. Therefore we decide to count an absolute values to be sure that none of the emotions is missing in the collective emotion.

2.2.5 Emotion Mining

A computational study focusing on research of deriving opinions, sentiments and emotions expressed in text is called emotion mining. (Liu, 2010; Kim and Hovy, 2004)
People express their feelings and sentiments about objects in many ways. One of these is the written form. In this form people use neither mimics nor voice intonation. The only medium they can use is the content of the text itself.

When we look at some text we can see that it consists of facts, opinions, orders, questions, etc. But interesting parts of text (for further evaluation) can be broadly categorized into two main types: facts and opinions. Both facts and opinions are expressions about object entities, events and their properties. The difference is that facts are objective and opinions are usually subjective. If we look at some opinion text we can easily evaluate author's sentiments by ourselves. As the concept of opinion is very broad, we will reduce them just to opinions which elicit positive or negative sentiments.

The Sentiment Mining Problem

The problem of sentiment analysis is quite complex. Many psychologists have claimed that certain emotions are more basic than others. For example (Parrott, 2001) claims that people have 6 types of primary emotions, i.e., *love*, *joy*, *surprise*, *anger*, *sadness* and *fear.* Each of these emotions can have also different intensities. There are a lot of sentiment words of how such emotion intensities can be expressed. The strengths of opinions are closely related to the intensities of certain emotions, e.g., *joy* and *anger*. However, the concepts of emotions and opinions are not equivalent although they have a large intersection.

When discussing subjective feelings of emotions or opinions, it is useful to distinguish two different notions: people's mental states (or feelings) and language expressions used to describe the mental states. Although we will take into account just 6 types of basic emotions (Parrott, 2001), there is also large number of language expressions that can be used to express them. Similarly, there are also a large (seemly unlimited) number of opinion expressions that describe positive or negative sentiments. Sentiment analysis or opinion mining essentially tries to infer people's sentiments based on their language expressions.

How to find a sentiment orientation of a word?

Most of the approaches used in sentiment analysis are based on evaluating sentiments which are expressed by polar word. Opinion words, which are also known as polar words, sentiment words, opinion-bearing words, are parsed from a text and then processed. E.g positive words are: good, nice, beautiful, flexible, responsible, etc; negative: silly, boring, etc. In our common language we use also opinion phrases and idioms. One example of such phrase is: "cost someone an arm and a leg". As you see finding examples of such words is quite easy work. But creating such lists manually can be really time consuming and also difficult to find most of the words. If we get all this words opinions expressions (polar words, opinion phrases) in once, it will create an *Opinion Lexicon.* Briefly, the most important words are adjective. One of the simple approach is to simply count polar words in the text. For the automatic creation *Opinion Lexicon* there are two approaches.

Dictionary Approach

For dictionary purpose two lists of polar words (positive and negative) are needed. The basic idea is to create some kind of seed lists for both categories, starting manually by writing down positive and negative words (the size of one list can be about 20-30 items). The main idea (assumption) is that for an opinion word ("good" from positive category) there exits it's antonym. We consider that this antonym word is suitable for the opposite category of the first word ("bad" for negative category). Also if we take a word from one category (positive) and we find its synonym, we can assume that this new synonym word will belong to the same category, too.

There are few online lexicons that can help automatically find antonyms and synonyms, e.g. WordNet (Miller, 2010).

However, not all synonyms and antonyms could be used: as some of them have different sentiment that they suppose to have by their group. In addition, some common words such as "great", "strong", "take", and "get" could occur many times in both positive and negative categories. This indicated that they need to develop a measure of strength of sentiment polarity (the alternative was simply to discard such ambiguous words) for determining how strongly a word is positive and also how strongly it is negative. This enables to discard sentiment-ambiguous words but retain those with strengths over some threshold (Kim and Hovy, 2004).

Corpus-based approach

Unfortunately, lexicons and dictionaries like WordNet do not include semantic orientation information. Corpus-based approach presents a method that automatically retrieves semantic orientation using indirect information collected from large corpus. The methods in the corpus-based approach are based on syntactic or co-occurrence patterns and also a seed list of opinion words to find other opinion words in a large corpus. (Liu, 2010)

One of these ideas was proposed by (Hatzivassiloglou, 1997). The approach relies on the analysis of textual corpora that correlates linguistic features, or indicators, with semantic orientation. While there are no direct indicators of positive or negative semantic orientation, the **“and” conjunctions** between adjectives provide indirect information about orientation. The assume is that the most of connectives with conjoined adjectives are usually of the same orientation. E.g. “fair and legitimate” belongs to positive category and “corrupt and brutal” to negative one. But parts of that conjunctions are in the same opinion polarity. Another example is is with **“but”** conjunctions, which usually connects two adjectives with different orientations. E.g. “this car is beautiful but difficult to drive”. This approach strongly depends on the corpus on which evaluation is made.

With Opinion Lexicon we are able determine the sentiment polarity of the word. By counting the number of sentimental words (also emoticons) in a message we can estimate the overall sentiment polarity of the sentence too.

2.2.6 Tools

For the purposes of this work we need a way to extract emotions (or emotional values) from text. By other words, we need a sentiment miming tool (as described in Chapter 2.2.5). For those purposes we use two different sentiment mining tools by which the datasets are preprocessed: SentiStrength (Thelwall, 2010) and ANEW (Bradley and Lang, 1999).

SentiStrength

SentiStrength is an algorithm invented by Mike Thelwall, Kevan Buckley, Georgios Paltoglou and Di Cai to extract sentiment strength from informal English text. (Thelwall, 2012) claims that previous sentiment detection algorithms tended to be commercially oriented and were designed to identify opinions about products rather than user behaviors. SentiStrength tries to fill the gap by using new methods to exploit the de-facto grammars and spelling styles of cyberspace. SentiStength applied to MySpace[5] comments and with a lookup table of term sentiment strengths optimized by machine learning, was able to predict positive emotion with 60.6% accuracy and negative emotion with 72.8% accuracy, both based upon strength scales of 1-5. (Thelwall, 2010)

5 http://www.myspace.com/Help/AboutUs?pm_cmp=ed_footer

ANEW

The ANEW is an abbreviation of for The Affective Norms for English Words. It was developed to provide a set of emotional ratings (values) for a large number of words in the English language. The ANEW goal is to to develop a 3-dimensional model rating in terms of pleasure, arousal and dominance, as the ANEW authors assume that the emotion can be defined as a coincidence of values on a number of different strategic emotions. (Bradley and Lang, 1999)

2.3 Previous Work

Modeling of collective emotions in online communities is a new research interest which started in recent years. Before the modeling of communities, the research focused on modeling emotions of a single individual; especially, the emotions of synthetic agents e.g. (Grach and Marsella, 2004; Rousseau, 1998). An example of multi-agent based approach to model human personality was done by (Doce, 2012; McRorie 2009). However, these approaches aimed to model a single agent/individual by the multi-agent system. Our goal is to model a community (group of individuals) where a single individual is represented by a single agent. Significant work was done in this area by a group of people from the CYBEREMOTIONS project (CyberEmotions, 2012), namely in the publications of: (Schweitzer and Garcia, 2010; Rank, 2010; Garas, et al. 2012).

As we mentioned in the Introduction chapter, CyberEmotions is a research consortium that observes and analyzes the Internet communication phenomena (text, sound, visual or a combination) related to emotional processes. (CyberEmotions, 2012).

One of the CYBEREMOTIONS' publications (Chmiel, et al. 2011b) observes *Collective Emotions Online* and *Their Influence on Community Life.* The work's results show that *„collective emotional states can be created and modulated via Internet communication and that emotional expressiveness is the fuel that sustains some e-communities.“* Using the automatic sentiment detection tool, *SentiStrength*, and stochastic methods applied on more than 4 million internet comments (BBC[6], blogs, Digg[7]) they show that *„Internet users' messages correlate at the simplest emotional level: positive, negative or neutral messages tend to provoke similar responses“.*(Chmiel, et al. 2011b)

As we mentioned above, significant work has been done in case of multi-agent based modeling of collective emotions. One of it is (Schweitzer and Garcia, 2010), where a modeling framework for research on the emergence of collective emotions is described. We can use that as a reference to determine some parameter values, as well as for comparing results of our model regarding collective states of whole groups, i.e. collective emotion.

The current model, however, is considered with modeling the emotional trajectories of the single agents in a group, i.e. the evolution of the internal states of individual agents, i.e. as a reference to other models' valence and arousal which represent emotional states of agents, but extending it with expectations and other cognitive elements.

(Schweitzer, 2007) claims that the essence of many phenomena do not depend on a big amount of their details. He accumulated a wealth of examples that demonstrated how simple components with simple interaction rules can arise into complex emergent behaviors. The reason is that while the collective emergent behaviors can emerge from the interaction of the features (parts) of complex systems, not all the features are relevant for arising behaviors. Just the interaction between few of them can be essential for emerging collective behaviors. (Schweitzer, 2007) also shows that a system of Brownian Agents holds these conditions of a model consisting from simple components (agents) and simple interactions rules. And even though Brownian agents are quite simple and do not directly correspond to the complexity of agents in MAS (see above), they can exhibit quite complex behaviors through their direct and indirect interactions with each other.

This Brownian agents approach has been used as the basis a model framework intended to model the emergence of collective emotions. This Agent-based model framework of collective emotions was proposed by Schweitzer and Garcia (Schweitzer and Garcia, 2010) and serves as a reference point for the modeling approach presented in this thesis.

6 http://www.bbc.co.uk/blogs/

7 http://about.digg.com/

3 Agent-based Model

Why an agent based model?

We choose a multi-agent based system to model emotional online users interactions as an abstract network of nodes connected by message exchanges. As we work with IRC logs, we take an IRC network as an example. An IRC network consists of users and channels. Channels (or rooms) are discussion boards where connected users can publicly communicate among themselves. If a user writes a message on a channel, every member of this channel can see this message and is potentially influenced by it. A user can be connected to more than one channel, to take part in more discussion groups at once. Additionally, two users can communicate between themselves privately, so that their messages are not visible in the public channel.

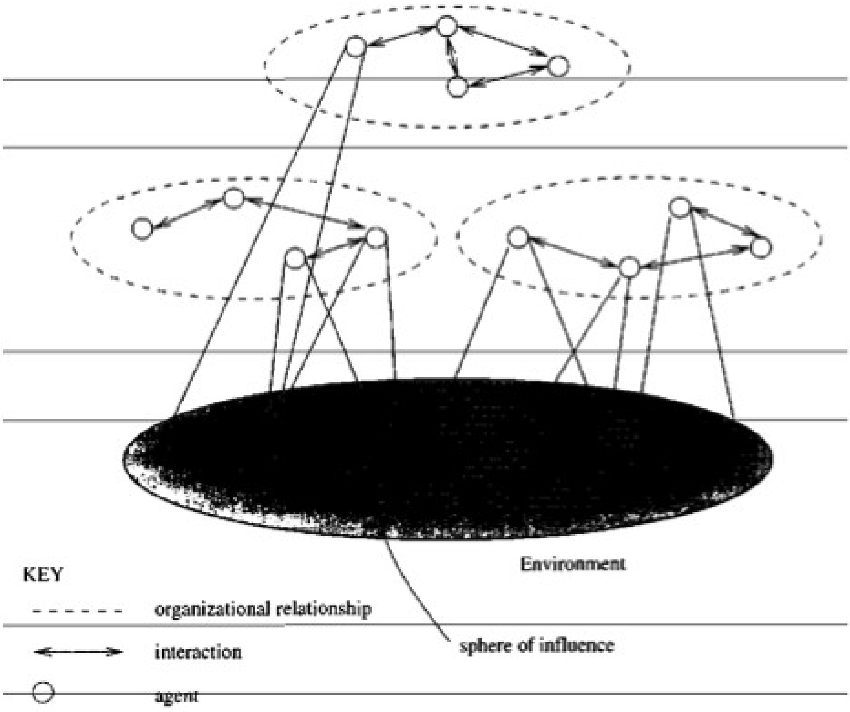

Figure 4. Typical structure of a multi-agent system. taken from (Wooldridge, 2002)

Figure 4 represents the typical structure of a multi-agent system. The graph-structure provides a suitable fit for the topology of users on IRC channels. The environment represents a channel, where all users are connected. An IRC user is represented by an agent. Interactions between agents represent messages that are sent between agents. As some users communicate by private messages, the communication is not going through the environment, so the agents are not influenced by it. Organizational relationships of agents can be mapped to different “groups” of users, e.g. junior users or senior ones on the one hand, or special types of user behavior, e.g. introverts, extroverts, aggressive, nervous, on the other.

3.1 Experimental procedure

The main experiment procedure consists of three parts: *extraction*, *prediction*, *evaluation* Before we start to model an online community, we need to create a multi-agent system model for it, the creation of a model is done by ***extraction*** of important information and features from an input data. The input data has to be separated into two parts. The first, and the major part is a learning data (in our case the first 70% of available data). This learning data are used to initialize and fill the multi-agent system model with agents and their personalities. The second and the smaller part is a testing data (i.e. the last 30%). After extracting (initializing) the model from the learning data, we are able to run this MAS and let it make predictions. The agents in the MAS are inactive and are in the „feeding“ mode, so they do not make any actions in the environment. Only agent's personality table and it's internal emotional state is updated by the agent *Update function* (See Chapter 3.2).

The ***prediction*** phase assumes that there is an already created model of agents of an IRC channel. However, it is able to work without any previous data about the channel. During the run, the model is awaiting any event (message) from the environment (channel). When the event occurs in the environment, the model fires messages to all active[8] agents. Here, the prediction starts. The first step of an agent is to decide whether the message was aimed at the agent (or not) and whether the message should be taken as a trigger and elicit an action in the agent. When the agent decides to predict its own reaction on the trigger, it still needs to predict the timing and the emotional values of this future message. Emotional

8 *active* agent is an agent whose last message in an environment is not older than some time constant (e.g. 30 minutes, one hour, three hours, ...)

value of this message is defined by its valence and arousal. As the *valence* of a message is defined on the interval on <-5,5> (integer – 11 values) and the *arousal*, too. The agent has to predict the emotional value from 121 (11x11) options.

One should keep in mind, that during the prediction phase, the model updates itself and its agents immediately by the *Update function* after a new event in the environment appears. The agents personalities are not updated by the values of predicted events, but only by the real values from the environment.

The last, **evaluation** phase of the experimental procedure checks the consistency between the predicted events and those that really happened in the environment. For every predicted event, the *evaluation* tries to find a real similar emotional event in the environment which happened in the similar time period. For deeper description of evaluation, see Chapter 3.5.

3.2 Agent

Generally, agents communicate to achieve their goals in the society (system) in which they exist (Huhns, 1999).

Goals & Motivation

In general, agents goals can be explicitly defined or not. As we do not use any semantic processing of the data that we use, so we are not able to explicitly determine what are the exact goals for individuals agents. By looking on any user that spends his time on IRC channel, we could see his goals as: communication about interesting topics with other people, finding new friends, or just trying to not be bored. However, all users have some motivation to be "online" and chat, otherwise they would not do that. In our model, the goals and motivations are derived solely from the activity of an agent on the channel.

For the purpose of modeling, we consider two types of agent action: joining/leaving an environment (channel) and sending a message to it. The central feature for the model is the the second type of action. An agent can send a message with 3 parameters: receiver, valence, arousal. A receiver of a message does not need to be included every time, it defines an agent at whom the message is aimed. A message can also be without receiver, i.e. be aimed at all agents in a channel. The values for valence and arousal describe the affective content of an agent's message, independent of the cause for the message which could be an external event or an internal event in the channel.

Thus, agents' external actions can be defined formally as the following tuple:

$$A = \{\delta, a, v\}$$

where δ defines the receiver of a message (which can be also undefined)

$$\delta \in Agents \cup \{null\}$$

a defines the arousal value of a message

$$a \in [-5 \dots 5] \wedge a \in N$$

v defines the valence value of a message

$$v \in [-5 \dots 5] \wedge v \in N$$

An agent operates in an environment, ***E***. The environment in which agents operate can be seen as a log (or history) of all agents' actions. Formally, we define:

$$\boldsymbol{E = e_0, e_1, \dots , e_n}$$

where ***E*** is the environment, and e_i is a message event:

$$\boldsymbol{e_i = T_i \delta_i Ac_i,}$$

where T_i is a timestamp of the external action, δ_i is its author and A_i is the content of the external action expressed by this agent at that time. The index *i* ranges from 0 to N, the number of actions seen so far. An environment is, thus, an ordered list of all external actions from all agents. **r**.

A *run* of an agent in an environment is thus a sequence of self-dependent pairs of environment states (e_i) and agent's actions (A_i):

$$\boldsymbol{run: e_0 A_0, e_1 A_1, e_2, A_2, \dots}$$

For our purposes we need an agent with internal state. To be able to better imagine an abstract definition of an *agent with state,* we can look at the following Figure 5:

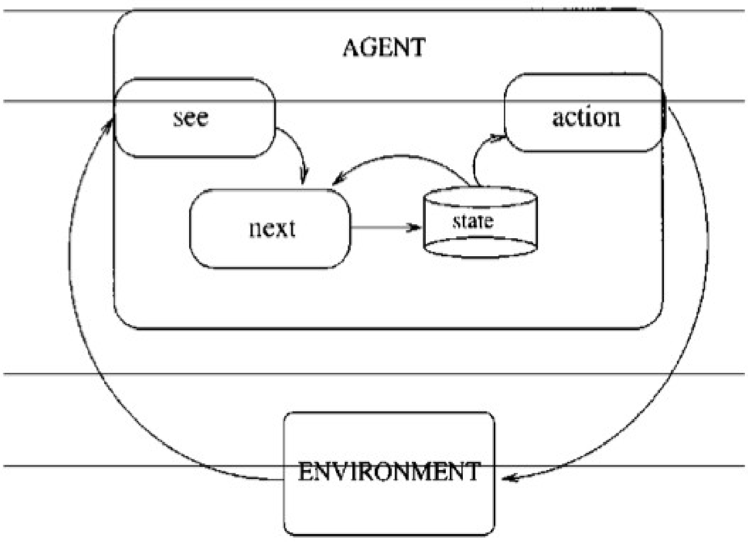

Figure 5. Agent that maintain state taken from (Wooldridge, 2002)

The action-selection function *action* is denoted as a mapping action:

$$action: I \rightarrow A$$

from internal states to actions. An additional function *next* is introduced, which maps an internal state and percept to an internal state:

$$next: I \times Per \rightarrow I$$

The agent that we define as a basic unit in our multi-agent system is an instance of a general model of *agent with state.* The internal state of our agent consists of a current emotional state as as a “personality table“:

$$I = \{\varepsilon, \Pi\}$$

An agent's **internal emotional state**, ***ε,*** where:

$$\epsilon \in [-5 \ldots 5] \wedge \epsilon \in R$$

The personality table, **Π,** is an unordered list (a set) of rules:

$$\Pi = [r_0, r_1, r_2, \ldots, r_n]$$

Where rule, ***r,*** is a n-tuple consisting of information about one more or trigger events that elicited a type of action in the history of the agent, and the agent's reaction to this trigger:

$$\mathbf{r} = T_t, \varepsilon_t, \delta_t, a_t, v_t, T_r, \delta_r, a_r, v_r$$

In folk psychological term, one can say that **ε** represents how an agent "feel" at a local state. (for more information about it, see Chapter 3.2.2).

The time units, T_t , T_r, are exact timestamps of trigger's and reaction's occurrences. The trigger parameters: T_t, δ_t, a, v_t represent a situation that occurred in an environment. Thus a trigger event is represented by it's author, arousal of a message and valence of a message. The reaction parameters: T_r, a_r, v_r namely: reaction time, reaction arousal, reaction valence.

Thus, a personality table can be seen as a selection of potentially abstracted, cases available for future decisions similar to case-based reasoning, while the rules themselves are derived from previous experience in the reference data similar to rule induction in machine learning applications.

Agent initialization

Before a run of an agent's program during the initialization (extraction) phase of a model, we we derive agents' personality tables by an observation on available data. However, not all agents appear in data during the initialization. During the prediction (simulation) phase, a model encounters new events and new agents that were not present before. One feature of the model is to be able to build-up a new agent with a personality table during the model run. This feature is provided by an Update function (See: Agent's UPDATE function), which is able to deal with both types of personality tables: empty or already filled ones.

Agent's UPDATE function

When an agent reacts with an *action Ac* on a trigger in an environment a *rule r* for a *personality table* ***Π*** is created as defined above.

If the agent's personality table ***Π*** already contains a rule r_1 which is sufficiently similar to *r,* we increase the number of occurrences of this rule by 1. Two rules are *same* when

$$r\ (\varepsilon_t, \delta_t, \varphi_t, \delta_r, a_r, v_r) = \mathrm{r}_1\ (\varepsilon_t, \delta_t, \varphi_t, \delta_r, a_r, v_r)$$

and

$$\left| \, |r(T_l) - r(T_r)| - |r_1(T_l) - r_1(T_r)| \, \right| < TIME\_EPSILON,$$

where TIME_EPSILON is a number of seconds, e.g. 10, 60, 300, etc.

If rule *r* does not exist in the agent's personality table Π, then the rule is added into the table:

$$\Pi = \Pi \cup \{r\}.$$

Agent's workflow

After we formally defined the agent, its initialization and the function of updating its internal state, we can summarize its workflow during the prediction (simulation) phase into these steps:

1. agent starts in an **initial state**, i_0. In our case, the initial states of agents differ for every agent. There are two main variants of an agent initialization. The first case represents a situation in which there is no previous information about the agent, $i_0 = \{\}$. In the second case we were observing the agent in the environment, so we are able to fill the personality table for this new agent, hence $i_0 = \{r_1, r_2, \dots\}$.
2. during a lifetime, an agent observes an environment and generates a percept ***see****(e)*. Right after that, the internal state is updated by the ***next*** function, $next(i_0, see(e))$. In our case, the *next* function filters new events from the environment and takes into account just those that are related to the agent and the receiver of an event is defined. By updating the *state* of an agent, the ***action module*** may perform a suitable action.
3. after the 2nd step of this workflow, an agent waits for any other input from an environment and the cycle repeats.

Environment

To tell whether the environment is deterministic or indeterministic is related to the MAS settings that we use for the modeling. By definition, an IRC channel communication is indeterministic, as the responses from the user are not only channel-related, but could be also influenced by external factors, i.e. talking simultaneously with a real person, reading newspapers, etc. A good example, how to simulate an indeterministic behavior of agent is to implement a probabilistic feature which cooperates with agent's processor for decision making. However, not all models that we use have to be indeterministic. It depends on a specific settings for every model, that we define in the beginning of modeling.

The environment in which agents operate is a discrete one. The number of possible actions that could occur in the environment, in the sense of discretized arousal and valence values, can be easily determined.

Emotional State

People behave differently if they are in different emotional states. Therefore, one of the main features of this model, is that it is able to determine emotional states of an agent. The interpretation agents' emotional states and personality represent agents' reactions on events that occurred in the past. It also represents the current emotional value of an agent, its attitudes, future plans and goals.

In other words, a good reason why we considered an agent emotional state is from (Gratch and Marsella, 2004):

“Moods (emotional state) *are an affective phenomena closely related to emotion. Typically, mood is distinguished from emotion as being more global, diffuse and longer-lasting. Moods are not “clearly related to a single object or piece of business in an adaptational encounter, as is the case in acute anger or fear” (Lazarus , 1991). Moods are important to model because they have been shown to impact a range of cognitive, perceptual and behavioral processes, such as memory recall (mood-congruent recall), learning, psychological disorders (depression) and decision-making.”* (Gratch and Marsella, 2004)

Going through a history of an agent communication we can observe agent's behavior on different events – messages (new one, already known one) in different emotional states. This behavior of an agent in a variety of situations can be used to approximate an agent's coping strategy in such situations. By the observation we can define an agent personality. Here, an agent personality is represented as an “n-dimensional table”, where the dimensions of the *table* are conditions (settings) of a trigger of a coping strategy of agent in a specific situation. One of the goals of this work is to find “suitable dimensions” of the table (personality). A priori we can assume that the necessary dimensions are “current emotional states”, “topics” and “emotional values of messages”. Possible dimensions could be “an author of a message” and “timing”. We also have to consider an agent personality in case of new topics of messages.

We take into account that, the personality of an agent is not stable, but it changes over a time. The changes are detected by the observation of an agent behavior. According to the attitudes theory *“Attitudes as context-sensitive construction on the spot”* by (Schwarz and Bohner, 2001), we specify a trigger to a specific situations in the personality of an agent as an *attitude.*(Schwarz and Bohner, 2001) also claims that *“we may expect a close relationship between attitudes and behavior only under some specific, and relatively narrow, conditions”*, which holds our definition of the agent personality. This theory also says that similar judgments are to be expected when people form similar mental representations of the attitude object and a relevant standard at different points in time. Here, “judgments” can be noticed as a “selection of a coping strategy”. Consequently, we can make a prediction of next agent's coping strategy in the each step of an agent communication.

What coping strategies do agents have?

The only thing we can measure in agents' behaviors is their chat communication, from which we can extract valence, arousal, dominance and the timing of agents' responses. This response represents an external indicator of the agent's coping strategy. We also have to consider the internal variable of an agent which is represented by his current emotional state. According to this, there are just few types of possible agents' coping strategies:

- no response – an agent does not respond (express) to a stimulus. The coping strategy consists just from a change of an agent internal emotional state.
- agent responses – additionally to an agent internal emotional state, the coping strategy consists of external signs (arousal, valence, dominance) of different intensities. Here, coping strategies can differ in intensities of the expression

3.2.1 Emotional trajectories

Above, we have defined agent *emotional state, personality (attitudes, coping strategies) w*hich represents an agent behavior during a communication in an online chat-like community. While we were defining attitudes, we said that attitudes can be measured over time. Through that, these attitudes measurements can be used to construct a personality table. With this personality table we are able to model an agent's coping strategy in the next step of his communication.

A coping strategy consists of 5 variables: emotional state, arousal, valence, author and timing. The values (intensities) of the first three parameters through time can be represented by curves in a two-dimensional chat-like diagram, where the first dimension (X) is time and the second dimension (Y) are the values of the variables of a coping strategy. By other words, the change of emotional state, arousal and valence over time are the emotional trajectories, which are represented by the curves.

As we create our model (table) of an agent personality, which is *de facto* a list of coping strategies (rules), we should be to determine a next coping strategy in a next time step by the actual events (settings) and agent's emotional state in the environment. This means that by taking the exact values for arousal, valence and a similar distance in time of an actual event in an environment, the model can extract a proper coping strategy from it's personality table for a next time step. This extraction of a next coping strategy represents our “prediction”.

3.2.2 Internal Emotional State of Agent

If we look into the all types of information that we have about an agent we can see that there is not a single explicit information about any of agent's internal states. Therefore, the only possible way how can we extract any piece of information about agent's internal state is to look further into the agent's expressions and use it for describing a simple model of an agent's internal state.

As we work with the terms of valence and arousal across the whole work, we will use them also here for determining an agent state. Anyways, the internal state is not represented just by the actual values of valence and arousal that were calculated from the last agent actions, but the values are additionally processed. Otherwise, it would not be an internal state, but an external state. On the other hand, the usage of "pure" not-processed values of valence and arousal to determine an agent state could be an interesting experiment.

Let say that agent's expressions directly correlate with his internal feelings. Therefore, we take expressions values as values of an emotion that was elicited in an agent in a local time. These emotional values have to be added to an agent overall emotional state. In Chapter 2.2.3 we described a way how emotion intensities can be combined by a sigmoidal or logarithmic function. We use one of these mathematical functions as a basis for the calculation of an overall emotional state of agent. The thing is that we always add a local value to the overall value by using this equation. Through it, an internal state of an agent will be represented by the overall emotional value. With such a model we are able to determine an internal agent emotional state in any time. Additionally, we can can add a decay functionality to it, so in every timestamp that an agent is silent, the values of his internal state will be subtracted by constant, so after a longer period of time, the agent will relax into a neutral emotional state.

One can doubts that the technique of the calculation of an internal state is too simple. Nevertheless, the one of the goals of this work was to develop a working model which is not too complex.

Valence & Arousal

Considering only an ordinal interval <-5, 5> for both valence and arousal we get 121 *(11 x 11 = 121)* possible agent's internal emotional states. Therefore, one can think about reducing this big amount of states into the concrete discrete emotional states (basic emotions) defined by it's position in the graph (see Figure 1. Alternative dimensional structures of the semantic space for emotions (Scherer, 2005)). The reduction can be done by defining a constant natural number (e.g. 1, 2, …) which represents an area around the specific value of valence & arousal. This area then represents a specific emotion.

For the purposes of internal emotional state, we decided to use the same mechanism of accumulation of emotional values as described in the *Combination of Emotions* Chapter (Chapter 2.2.3). We assume that the emotions accumulates in an agent in the same way they are accumulating for the combination of emotions.

There we base the emotional accumulation in the internal state of agent on a logarithmic combination (Reilly, 2006):

$$f_{v,a}(em)=0.1\times\log_2(\sum_{em} 2^{10\times intensity(em)})$$

Figure 6. internal emotional state accumulation

where f_v represents the valence of the current emotional state and f_a represents the arousal of current emotional state. Therefore we have defined the agent internal emotional state as a pair of real number *(v, a)*.

Equilibrium state

We assume that an agent (or human) does not remain in the one emotional until a new emotional event comes. When an agent does not react for a longer period of time the model decay it's internal emotional state into the equilibrium (0) values. Therefore, for every single timestamp when agent does not react the model increase (or decrease) the emotional value by a constant ***ED*** with respect to the positiveness or negativeness of the internal emotional state numerical value, what means that the *arousal* and *valence* of the internal state are influenced by the decay at the same time

3.2.3 Personality Table

"Personality traits reflect individual differences in reactivity to emotional and affectively valenced environmental cues." [Revelle2010]

In our modeling, we represent the agent's personality by the *personality table.* The personality table could be imagined as a list of rules (coping strategies). This list describes how an agent behaves in different types of situations. The personality table does not contain any other data than these behavioral rules.

Behavioral rule

The behavioral rule holds an information about trigger(s) that occurred in the online environment and somehow influenced the agent, and it also holds an information about the agent's reaction on that situation. The first mentioned information represents inputs (triggers) and the second information represents outputs (reactions) for the behavioral rule. All the triggers and the reactions are basically the same type of a data, but with different values. We define this type of data as *Emotional Event* (see below). The behavioral rule also holds information about an agent internal emotional state (see Chapter 3.2) in the time of the agent reaction on the inputs.

Note that the behavioral rule corresponds in psychological theory to Coping (see Chapter 2.2.2).

According to Chapter 3.2.2, we estimated an agent's internal state in any time. The estimation of the agent's internal state directly depends on agent's reactions on a concrete activity in the online environment. From the previous observations we determined the basic rules, because the events from which we determined these basic rule were clear in case of: one agent's reaction is directly related to one “foreign” agent action. So the stimulus (trigger) of the agent's reaction was easy to find. These basic rules are the build stones for the agents personalities.

After all fully-described messages are processed, we use the gained knowledge about the agents to explain messages, for which we know recipient, but there were more messages as a respond to a situation. Such a bunch of messages we would call a “complex trigger”. Complex trigger is a list of messages (partial) triggers that elicited an action (emotion) in an agent.

For every complex trigger we try to process all the complex triggers by disassembling the complex trigger to basic rules. If there was enough known basic rules (extracted from personality tables) that were similar to the disassembled rules, we explained new basic rules. We added these new basis rules into the agents' personality tables.

Emotional Event

The emotional event is a basic stone on which we build the whole agent personality and *de facto* a basic stone for the whole representation of our multi agent system. Emotional Event is any message (event) that occurred in the online environment. This message can be any text message from a user to other users or the information about the joining or leaving the environment by agents. We call this message as an environment event. The emotional event holds the timing of the environment event happening, it's author, the direct recipient of the environment event (is not always defined), the **valence** and **arousal** of the message and a type of the event.

The personality table and all its subparts represent an agent personality.

3.2.4 Agent's default personality

When a user comes into a chat room (we will call such a user a "*newbie*"), we do not have any previous information about his behavior or his personality. Additionally, when a newbie posts just few messages, there is not enough data (information) for an agent to model the newbie's personality. However, we should be able to model a newbie's behavior, as his future posts can significantly influence emotional states of most agents. E.g.: consider the reaction of users in online support chat room of a software application to an event such as a new user is claiming that the software has a big security issue. Also considering the goal of this model to application as online decision-support for interactive conversational systems, modeling default personalities is essential as newbies quit the support board after few questions. Hence, it is useful to set up a model of a default agent personality.

Here we propose two different perspectives for a default personality implementation: *several prototypes for all* and *one for all*. The first method is based on clustering all personalities in the groups and then create several typical behaviors specific in the online environment. The latter one is based on merging all personalities into one default personality with a specific procedure for selecting *rules* from it.

“Several prototypes for all“ default personality method

Because we do not have any information about newbie behavior, one thing that we can assume about his personality, is that there is a big probability that he will react on new messages (events) with intensive emotional valence of any topic *(=a message valence passes through an new valence threshold)* ; and on new messages (with any emotional valence) which topics are already included in newbie's post(s). On the other side, any further general information about newbies behaviors can be observed on a history of a communication of all agents in a chat room. Through a history we can see many newbies behaviors since they joined a chat until they leave. Here we can observe few typical behaviors of newbies when they join a discussion. E.g. *“how do they react to a new topic?”*, *“how do they react to old-users with different kinds of personalities”*, *“how can they get used to new users and their communication style in a new environment?”* and with respect to previous items: *“how are their emotional tables filled in the beginning of they participation on a discussion?”*. With this observation we are able to create few **typical newbies personalities**, which can be easily associate to a newbie according to his few posts in a discussion. This typical newbie personality corresponds to the part of a personality of “old users” which describes agent's behavior in case of new events.

Calculation of default personalities

After the extraction of all agents personalities from a learning data, we can group all these agents into similar behavioral groups. As a key for grouping we can use a number of similar rules in agents' personalities. Here we define a parameter ε_p, where:

ε_p in N *and* ε_p > 0

ε_p represents a number of similar rules that are needed to decide whether two agents should be in the same behavioral group or not.

For an estimation, one can take into account an average number of messages per user, but this parameter is set empirically and according to the data input data. Can be automatically counted as an average number of messages divided by some constant (e.g. 5, 7, ..). The parameter ε_p directly influences a number of behavioral groups created. If ε_p is too big, a number of groups is too small. And if ε_p is too small, a number of groups is too big.

When processing agent's personalities, we should not take into account all rules included in their personalities tables. Hence, we should look at first *NEWBIE* rules of all rules in one table, where:

NEWBIE in N and NEWBIE > 0.

Workflow of creating default personalities for a single agent:

- iterate through all groups and check whether the first agent (prototype) from a group has a number of similar rules greater than ε_p.
  - if holds, add an agent to the group
  - if not, create a new group and add the agent into it

This workflow is repeated for the all agents. After the all agents are processed, default personalities of agents are created

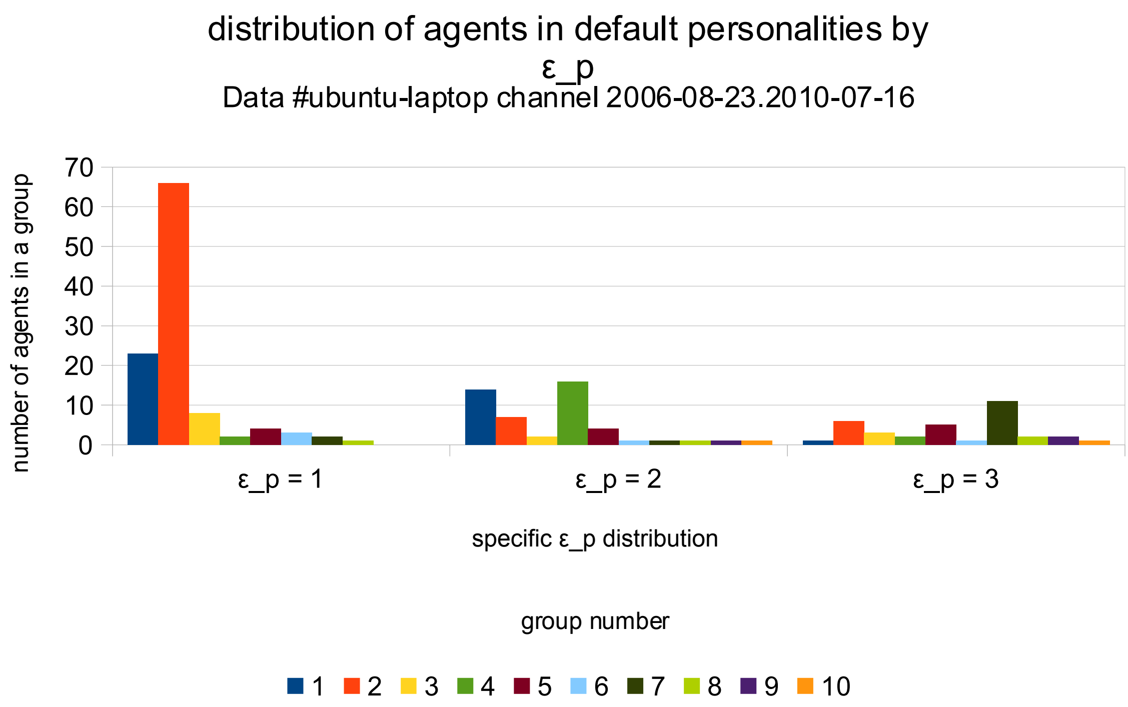

Figure 7. Distribution of agents in default personalities by ε_p

Usage

When a totally new user joins the channel we wait until he does reply on his first event. Then we know the first rule of a newbie's personality and we can find a most similar default personality. The newbie adopts rules from the default personalities until he not reach a limit (*NEWBIE* constant) of sent messages and do not “transform” from a newbie to an experienced agent.
While processing and predicting new emotional events, newbies with adapted personalities tables do not take an author (sender) in rule into account. Newbies typical personalities could be derived also from the far history when users that do not attend an environment anymore could be included into the rules. Such rules for default personalities would be useless if we check for authors when determining whether two events are similar or not.

“one for all” default personality method

The next idea of default personality model is more stochastic based than one before. Instead of creating several default personalities, we can create one „default personality“ including newbies' personality rules of all agents. As a start we create one empty default personality. Going through input data set, for every agent that is observed, we add his first X reactions to this one default personality. After all events (agents) are processed the default personality table is well-filled and ready to be used for a prediction. If a newbie agent has to react on some trigger, he will use this default personality as it's own. If there are more rules with the same trigger values in this personality, agent can use standard rule selection and choose a rule randomly.

To improve this idea by having a newbie agent to behave in some “types” or “groups” of personalities, we can take into account also agent's previous selections of rules. Therefore, for every agent rule selection (prediction) from the newbie personality that was correct, we can store an information about an author to which this rule was related to. For further rule selection, an agent will try to find a rule firstly from authors whose rules he already used. We call this method for using default personality as “**one for all**” method.

We have to describe one part of newbie personalities that is not based on rules from a personality table, because agent's react before sending any previous trigger message. When a totally new agent joins the online environment, it often happens that he writes a new message to the channel, e.g. to ask a question on irc support channels, or generally say hi to the channel. However not all newbies do that, so we can make a stochastic observation and calculate a probability of such action from the learning data. During the prediction phase of a model we can use this probability to predict whether an agent will write a first message into a channel or not.

Note that by this stochastic method we do not use any other specific rule to predict whether an agent will send a message or not. From the point of a classification view, we have a set of N and we are classifying them into two categories by a classifying function with a correctness of C (the function is random classifier).

The idea behind including this feature in the model is that there is a significant flow of new users in IRC channels that write their first messages into it. However, these very new messages can not be predicted by the model.

To calculate emotional values of these new predicted messages, we only need to take the weighted average of all “first messages” to get arousal and valence of such predicted messages.

3.3 Data

Generally, the input data that we use in the project are annotated IRC logs from Ubuntu's support channels. For those who didn't get into the touch with IRC yet, it's a good point to show an example of such IRC communication:

11:30:59 02-29-2012 Jane joined #ubuntu

11:32:01 02-29-2012 <Jane> Hi all! I have a problem with sound on my laptop running on the latest version of Ubuntu.

11:33:04 02-29-2012 <Kate> hi Jane! What kind of problem do you have?

11:33:35 02-29-2012 <Jane> I don't know, the music is just not playing.

11:34:57 02-29-2012 <Kate> Try to install the newest Windows version! ;-)

11:35:23 02-29-2012 <John> Kate, your solutions are always sooo meaniningful.

11:36:03 02-29-2012 <Kate> look who's talking...

11:37:30 02-29-2012 <Jane> any better idea? :-)

11:38:44 02-29-2012 <John> Jane, have you looked at your sound card driver? If not download and install the newest one.

11:39:04 02-29-2012 <John> Kate, at least i'm not telling bullsh*ts to newbies

11:39:21 02-29-2012 <Frederik> guys, you're always so funny when you start the argue. Hahaha

11:40:04 02-29-2012 <Kate> shut up, nobody's interested in your stupid stuff

11:42:03 02-29-2012 <Jane> hmm, it's starting go crazy here. i will try john's idea, but thanks to all!... bye

11:42:23 02-29-2012 *Jane left the channel*

11:33:55 02-29-2012 <John> that's it ;-)

11:35:01 02-29-2012 <Kate> at least she 's not bothering around anymore:-)

Text 2. IRC channel log example

In the IRC communication above (Text 2), you are able to see how such a communication on IRC could like. However, this was a very short and fictive example. The number of discussing users are often greater than 2 and also the dialogs are a bit longer.

For our purposes we use IRC logs of several ubuntu channels from the period of time 2006- -2011.

As already mentioned, the data that we use are not pure IRC logs, but they are preprocessed and annotated. The first preprocess step is anonymizing of the user names. The communication on IRC is private. Additionally, anonymizing tool convert IRC names into number Ids which is helpful for using this data in the multi-agent system.

The second part of preprocessing is a sentiment annotation of messages (s*ee Chapter 2.2.5 about emotion mining).* By this preprocessing, words and semantic information from the message is replaced by the information about sentiment (valence & arousal) value of a message. The preprocessed IRC log from (Text 3) could look like:

```
11:30:59 02-29-2012 11111 joined  #ubuntu
11:32:01 02-29-2012 <11111> receiver: None; arousal: 2; valence: -1
11:33:04 02-29-2012 <22222> receiver: 11111; arousal: 1: valence 0
…
```

Text 3. Processed IRC channel log example

Here you can see that every message in the communication is annotated by its receiver, arousal and valence.

3.3.1 Data structure

The input data that we use is an hdf5[9] database file. The database is a non-relational one and consists of 3 tables: *Nodes, Linkevents* and *LinkeventAnnotations.* Note that in the following, paragraphs we will only list the elements of the data that are used in the simulations rather than all the annotations present in the datasets.

Linkevent table

Column Name	Column type	Column description
start_time	32 bit timestamp	a start time of an event
end_time	32 bit timestamp	an end time of an event (mostly the same like start_time)
sender	32 bit integer	an ID of an author of an event; the value is not null
recipient	32 bit integer	A target user of an event, the value can be null

Table 2.Linkevent table of datasets

LinkeventAnnotation table

Column Name	Column type	Column description
valence	8 bit integer	valence evaluation of an event. Interval: <-5;5>
arousal	8 bit integer	Arousal evaluation of an event. Interval: <-5;5>

Table 3.LinkeventAnnotation table of datasets

9 http://www.hdfgroup.org/HDF5/ HDF5 is a data model, library, and file format for storing and managing data. It supports an unlimited variety of datatypes, and is designed for flexible and efficient I/O and for high volume and complex data.

Nodes table (users)

Column Name	**Column type**	**Column description**
Num	32 bit integer	Anonymized user id number
first_time	32 bit timestamp	first time used on this channel
last_time	32 bit timestamp	Last time used on this channel
node_type	Enum (user, bot, operator)	Defines whether a user is a *normal user, bot* or *channel operator*

Table 4. Nodes table (users) of datasets

3.3.2 Data characteristic

The data that we use as a reference is quite specific: the chats come from the Ubuntu IRC support channels. Users in many of those free support channel do not hold long discussions very often. Generally, the data suggests that typically a new user comes to a channel and asks a very specific question about a problem that he faces with Ubuntu at that time. Let us call this new user a *newbie*. After one or two messages from this newbie some other more-skilled user replies to his question. But what very often happens that no one reacts to this newbie and this newbie leaves the channel and will not come there any more.

In this section, we will compare two ubuntu support channels that have most significant difference in their characteristic, #ubuntu-laptop and #ubuntu-irc channels.

Here are few important statistic points about these channels:

	#ubuntu-laptop	***#ubuntu-irc***
Starting date of logging	*23-8-2006*	*16-1.2008*
Last date of logging	*16-7-2010*	*18-7-2010*
Normal messages	10860	39379
Empty messages	5	9
URLs	253	1588
Action	77	834
Nick changes	604	4088
Channel Joins	13793	0
Channel Lefts	1066	0
Kicks	0	0
Topic changes	99	12
Channel modes changes	3	0
Amount of users	1369	1080
Average number of messages per user	8,9	19,6
*Newbie action probability**	0.248	0.0
Total:	26760	45910

**see description in Chapter 3.2.4 (agent default personality)*

Table 5. Stats of #ubuntu-laptop and #ubuntu-irc channels

Table 5 shows the main quantitative differences between these two channels. #ubuntu-irc logging was done in almost 50% shorter period of time than #ubuntu-laptop was. On the other, number of *normal messages,* which is the most important information that we use to create our model, is almost 4 times greater in #ubuntu-irc than in #ubuntu-laptop. As we can see for the both channels, the information about channel joins and lefts is not consistent or missing. This is unavoidably caused by the logging process that the data is based on.

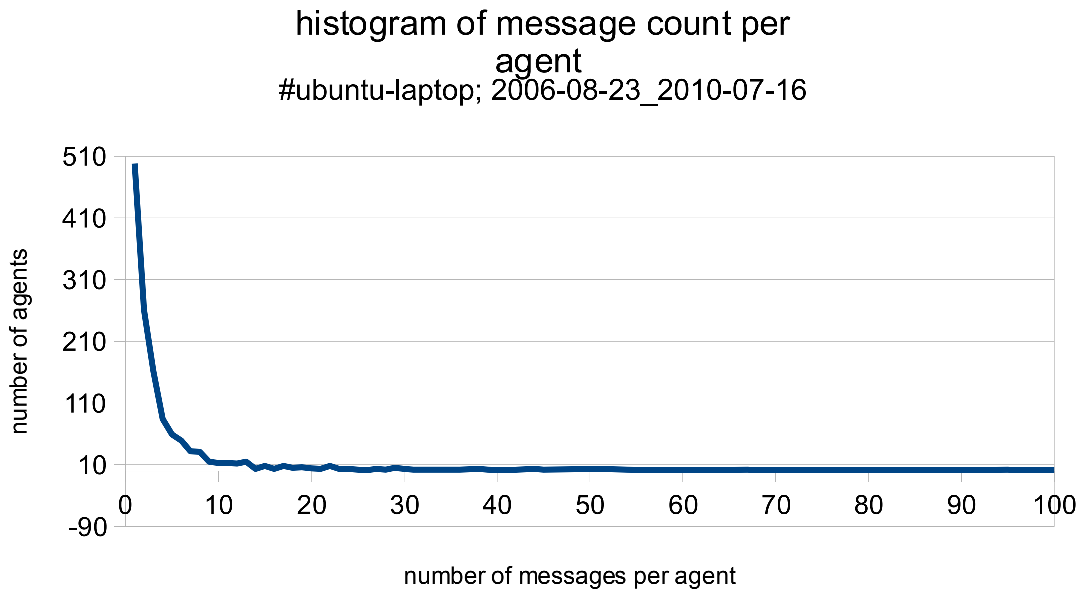

Figure 8. Histogram of message count per agent in the #Ubuntu-laptop channel

In the next figure (Figure 8) you can see a histogram of message count per agent in the #ubuntu-laptop channel. There are 498 users that sent only one message and 261 users that send two messages. These numbers support the characteristic above, that a lot of new users joined the channel, wrote few messages and left the channel forever. Such a user behavior is very hard to predict and without no information about external world we are not able to predict such random newbies visits at all. Anyway, if such random newbies start to join the channel in a very short period of time, one could deduce that something important happened around Ubuntu topic that forced new users to join and ask questions.

On the next Figure 9 (#ubuntu-irc), we can see two main different characteristics comparing it to Figure 8. #ubuntu-irc channel has 148 users with one message and 121 users with two messages per whole communication, what is about 30% less than in #ubuntu-laptop. We define a short-time users as users whose number of all messages in a channel is less than a constant *DefaultMax* (e.g. 3, 5, 6, ...). A ratio of users, that are not a short-time users, can be used to create default personalities. The ratio of long-time users according to number of all users is calculated by:

$$ratio = \frac{|agents_{long}|}{|agents|} \quad ,$$

where **|agents|** is a number of all users in a channel and **|agents$_{long}$|** is a number of all non short-time users (long-time) in a channel (users whose number of all messages is greater than DefaultMax).

For #ubuntu-laptop: ***Ratio_laptop*** *= (1369-(498+269)) / 1369 =* ***0,44.***

For #ubuntu-irc: ***Ratio_irc*** *= (1080-(148+121)) / 1080 =* ***0,75.***

This ratio is directly related to the effectiveness of our model, therefore a model simulation should be more precise on a channel-data with greater ratio of long-time users.

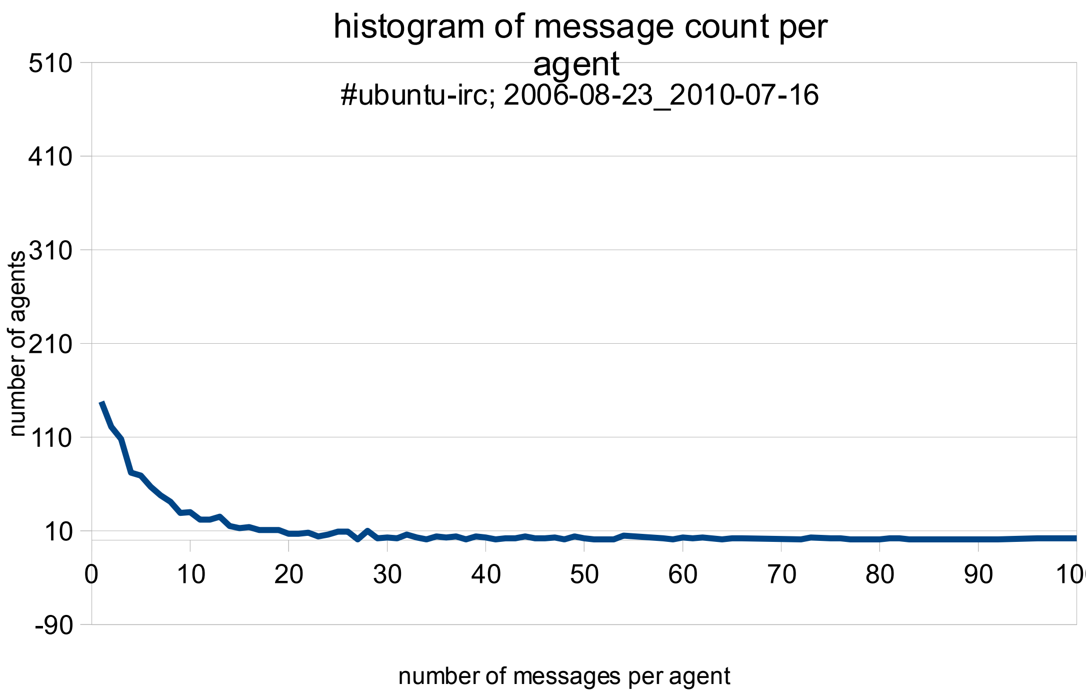

Figure 9. Histogram of message count per agent in the #Ubuntu-laptop channel

3.4 Model implementation

We implemented the model with the ***Python*** programming language[10] version 2.7. For our purposes we used Object Oriented Programming principles. Therefore we separated all the important modules (parts) of the model into the different classes. The main class is called "**MultiAgentSystem**" which handles all agents creation, initialization, workflow and all the communication between the environment and agents. The next important part of an MAS is an agent itself, which is implemented in the "**Agent**" class. Agent class processes all the input events, stores its and others agents' reactions and creates its own personality table which forms its behavior. The agent's personality table is represented by "**PersonalityTable"** class which is *de facto* a list of rules how the agent reacted on different events in the past. A rule is an emotional situation ("**EmotionalSituation**" class) which consists of trigger event parameters and also agent's reactions parameters. Every single event is defined by it's emotional value ("**EmotionalValue**" class) which holds the information about *arousal* and *valence.* The class also takes care about comparison of the similarity of two emotional values. Besides a personality table, an agent is also defined by an emotional state available for every time-window during which an agent is active. The emotional state and its workflow is performed in the **"EmotionalState"** class which can also hold multiple different implementations.

For the prediction and evaluation phases we created helper classes called: **"Predictor"** and **"Simulator"** which use an instance of the MAS. For the input & output purposes we created a **"Saver"** class. The input data is stored in the HDF5 format (See Footnote 9, page 10) which is very useful for working with big datasets that we used.

The main (error sensitive) parts of the code are covered by unit tests to help with a continuous increase in the quality of the code while avoiding regressions. For more information about the implementation, see the code attached to this thesis and the documentation inside it.

3.5 Evaluation

Before we specify what kind of features and performance of the model we want to

10 http://www.python.org/

evaluate, we need to define a measure of correctness of a single event prediction.

A prediction of a single event is deemed correct when an action predicted, A_p which is defined as:

$$A_p = \{T_p, \delta_p, a_p, v_p\}.$$

matches a real event A that appears in an environment: For a match the authors of the events must be equal. For every other value we define a maximum delta (Δ) by which a difference of the real and predicted values is considered to be correct or not.

T_p is a timestamp when the action should occur. The T_p is correct when:

$$\Delta_t = | T - T_p | \text{ and } \Delta_t < max\_\Delta_t,$$

where max_Δ_t is from N.

The predicted arousal and valence a_p, v_p are correct, when the distances of Δ_a, Δ_v values (on the possible intervals) between the real and predicted values are lower than max_Δ_a and max_Δ_v. max_Δ_a and max_Δ_v are from R.

The correctness of the predicted event holds when all the predicted event's parameters are correct according to the real event.

According to our goals and hypothesis we evaluate 3 types of results of the model:

- **Prediction of all single individuals**. For every single message from an agent in an online environment we check it was successfully predicted by the model or not.
- **Prediction of all newbies' behaviors.** Here we measure the correctness of the model for every actions (events) that were performed by newbies. Only the actions performed during users' newbie state are taken into account. The result of this measure will tell us more about the model suitability for online support-decision for interactive conversational systems.
- **Prediction of collective emotions.** Prediction of collective emotions leads us to group multiple emotional events are close together in mean of time they appeared in an environment. For this evaluation, we collect the predictions of the model over the whole prediction time period and calculate global values for total posts, average and variance of valence, average and variance of arousal, and a characterization of the occurrence of collective emotions in that set of posts by counting the number of

users expressing a similarly valued emotional state in a large (sliding) time window, similar to the working definition of collective emotions used in (Schweitzer and Garcia, 2010). These values are then compared for the actual data of the same time-period.

Note, that the emotional values that we use in the evaluation are rounded numbers of valence and arousal, therefore we allow small errors in interval of (-0.5, 0.5).

4 Results

In this chapter we provide an overview of the simulation results that we achieved with the model extracted from three channels taken from the reference dataset, for different time intervals. For the model configuration, we used two different settings: a strict and a less strict one. The strict settings allows no errors in emotional predictions, so the predictions of events must be *de facto* identical to the reference data. On the other hand, the second model setting is less strict and slightly more flexible for the emotional value variation, leading to more useful results.

The Multi-agent system was initialized by learning on data from the timespan between 28.8.2006 up to 13.4.2007. The running evaluation and prediction were run on the data from: 23-09-2006 to 7-10-2006.

We tested our model on 3 different parts of the dataset, extracted from the logs[11] of IRC communications on channels: #ubuntu-irc, #ubuntu-website, #ubuntu-laptop.

As you could see before in the Data Characteristic (Chapter 3.3.2), the number of new users, average number of messages per user or the frequency of events per time unit significantly differ between datasets.

For all the simulations we use the term of "timewindow" which represents the smallest time unit available for all the data and corresponds to 60 seconds.

We do several measurements for different periods of timewindows. For every simulation we compare the emotional trajectories (See Chapter 3.2.1) of the reference dataset events and the predicted events. Emotional trajectories are defined by the total valence and arousal. The total value is an accumulation of all absolute values (valence or arousal) for a specific timewindow.

Model Configurations

We use two different configurations (settings) of the MAS for the simulations. The configuration options of the MAS is defined by the Table 6. The configuration is divided into the two main parts: *Predictor params* and *Agent* params.

11 http://irclogs.ubuntu.com/

Agent params are used only during the extraction (observation) phase, when the model is created from the datasets. Namely:

Number Of Timewindows To Look Back For Trigger. During the extraction phase, we can see that an agent reacts on some events in an environment. However, to be able to find a trigger of that reaction, we need to lookup in the past to find out what was the trigger. This parameters sets the timewindow limit, how far back should the model look.

Personality table creation - Epsilon emotional. This is the parameter (which was not varied for the experiments) to be able to decide whether to emotional events are similar or not. The *emotional epsilon* is the maximum difference for valences and arousal to be considered as similar (See Chapter 3.2 - Agent's UPDATE function).

Emotional state decay. As we mentioned in Chapter 3.2.2, Agent's equilibrium state, we use a constant by which we multiply current emotional state values of an agent for a decay to equilibrium. As agent's does not react and the intensity of its internal state decreases, we use this parameter as a decay in every timewindow that agent does not react

The second dataset of configuration file is *Predictor params,* which only influence the behavior of the *prediction phase*.

Epsilon emotional parameter stands for deciding whether a predicted event was emotionally similar to a real event that happened in the environment.

Epsilon timing is the time interval around the predicted time, when the predicted event is considered as correct.

Epsilon timing – active user (seconds). During the prediction phase, some agents in the environment do not respond for a long time. If this long-time is overlapped, the user is not going to be triggered, until he does not make any reaction in the environment.

	Configuration 1	*Configuration 2*
Predictor params		
Epsilon emotional	0	1
Epsilon timing (seconds)	120	300
Epsilon timing – active user (seconds)	600	600
Agent		
Number Of Timewindows To Look Back For Trigger	20	20
Personality table creation - Epsilon emotional	1	1
Emotional state decay	0.99	0.99

Table 6. : Simulations' Configuration Settings

The first configuration (Configuration 1) is very strict in emotinal prediction, as it does not allow any kind of "error" in the prediction (Epsilon emotional = 0). However, as you can see later, the model with such a configuration gets quite impressive results in case of predicted emotional trajectories.

The second configuration (Configuration 2) allows "error" in an emotional prediction by 1. Also the *Epsilon timing* is set to 600 seconds which accepts reactions that are far away (by time distance) from the prediction. By setting the configuration file to less strict, the model should achieve better results.

Overall Simulation results – Configuration 1

Firstly, we extracted the model for some timewindow period (e.g <0, 3000>). Then we started the simulation on the period directly following the extraction window 10% of the size (3000, 3300>.

In Table 7, you can see a general overview of the model performance over all three datasets. The mentioned table is also a good example of how all the results overviews over all the simulations are presented . The structure of such a table is following:

Total events stands for all the events that happened during the simulation time period. The total events includes not only messages from users, but also non-emotional events such as: joining/leaving a channel, nickname change, etc.

Correctly predicted events is a number of all correctly predicted events during the simulation phase. Correctly predicted event is a virtual event which is created by the model and is similar to an event from the environment which happens in the predicted time period. For more details see Evaluation Chapter 3.5.

Incorrectly predicted events are events that were predicted by the model, but didn't happen in the environment. That means that there is no similar event to the predicted event found in the predicted time period.

Not predicted represents a number of all events that happened in the environment, but were not predicted by the model.

Events with recipients is a number more related to the characteristic of a data as a number characterizing the model performance. “Recipient” means that an event has explicitly defined a recipient of a message. This is important aspect for the data modeling we do. The basis of our model is to be able to derive an emotional trigger from the environment, to get an initial appraisal conditions for eliciting an event in a recipient agent. On the other hand, the model does its own heuristics of finding a recipient of a message. That's the reason why you can see in some overviews, that the number of events with recipients is less than number of correctly and incorrectly predicted events together.

Simulations of Configuration 1

Here we can see the overview results of the first simulations done on timewindows: 3000-3300, where the models were extracted from 0-3000 timewindows datasets and model was derived and predicted with the Configuration 1 settings

Events	#ubuntu-irc	#ubuntu-website	#ubuntu-laptop
Total:	808	689	1384
Correctly predicted:	52	106	0
Incorrectly predicted:	106	109	25
Not predicted:	756	583	1384
Events with recipients:	136	75	15

Table 7. Simulation stats for IRC channels. Model extracted from 0-3000. Simulations run on 3000-3300. ***Configuration 1***

From the table above, we can see that the model was mostly successful on #ubuntu-website dataset, then slightly worse performance was on #ubuntu-irc and with not even one correctly predicted event on #ubuntu-laptop dataset. Such results could be explained by the data characteristics of the datasets. For example, as you could see in #ubuntu-laptop data characteristic from Chapter 3.3.2, there is a very small amount of stable (old-experience) users and a lot of newbies with overall number of 1 message per whole time spent in the channel. As there were not many discussions (online communication) at all, the model was not able to derive personality tables properly and to apply them on new events.

By looking at the data characteristic of #ubuntu-irc channel, we can see that the majority of messages occurring in the channel comes from the old-experience users and the amount of newbie users and newbies messages is a minority. Here we can see the opposite situation as mentioned above about #ubuntu-laptop channel. As there are much more discussions and significantly greater average of number of messages per user, the model was successful during creation of users' personalities and henceforward applying them in the simulation. The analogical situation applies with the #ubuntu-website channel.

Comparing collective emotional trajectories -Configuration 1

First, we evaluates the results from the #ubuntu-irc data. On the figures (Figure 10, Figure 11), you can see the emotional trajectories of the #ubuntu-irc channel (original and predicted ones). You can see a comparison of emotional trajectories from the original data and also emotional trajectories that were predicted by the model. The predicted emotional trajectories contains correctly predicted events as well as incorrectly predicted events.

Every figure's graph consist of 4 parts (type of data) of emotional trajectories: arousal, valence, number of events, number of actions. The X-axis represents a timestamp[12] by which emotional events happened in the environment. The Y-axis represents a numerical value (count) of specific parts of emotional trajectories.

Count of arousal (or valence) represents the sum of all arousal (valence) absolute values (from all the events) in a specific timestamp. Number of events and number of agents (authors) in a specific type represent the total number of events that happened in the environment and total number their authors. In all the following figures in this chapter, you can see that total valence and arousal are very related values there is only slightly difference between their trajectories.

In the figures mentioned above you can see that there are 3 main peaks. The first significant peak is in the beginning of the graph of the original values. The similar peak, for the same timestamp, is included in the predicted graph, too. The same situations occurs on the peak near the timestamp 1210300000 and also in the end of the graph. The values of the original emotional trajectories are slightly greater. However, what is most important is that the most significant collective emotions, as indicated by the peaks, in the predicted events are consistent with the reference data.

However, for the better demonstration of similarities between emotional trajectories, we made a direct comparison of predicted events' valences (arousal) and original events' valences (arousals) in the Figures 12, 13. There you can directly see, that the predicted trajectories “imitate” the original trajectories in case of “bigger” peaks.

12 A **timestamp** is a sequence of characters or encoded information identifying when a certain event occurred, usually giving date and time of day. In the whole work we use Unix time, the number of seconds since 00:00:00 UTC on January 1, 1970

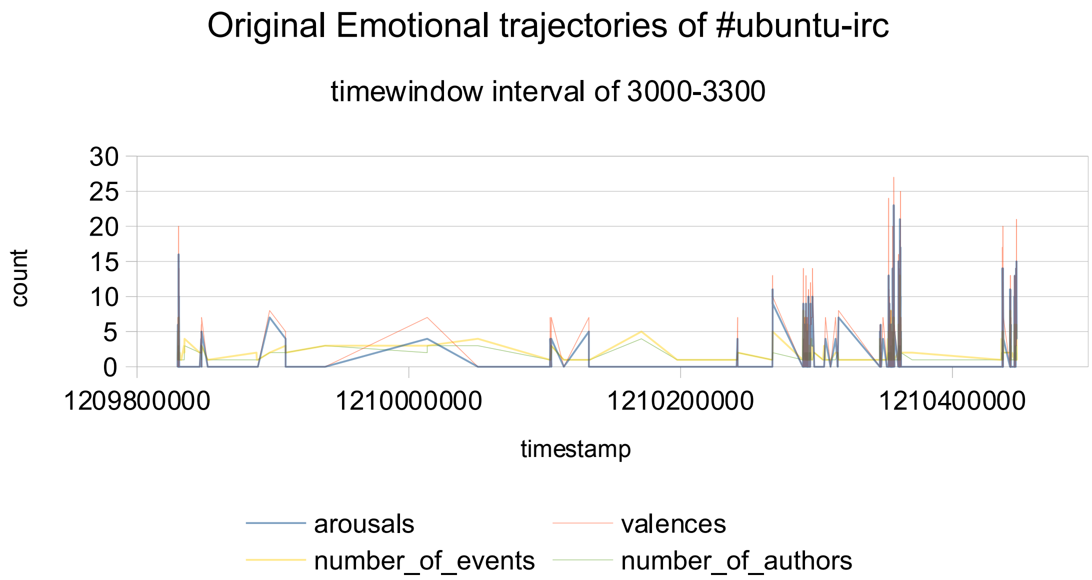

Figure 10. Original Emotional trajectories of #irc-ubuntu on timewindow interval of 3000-3000

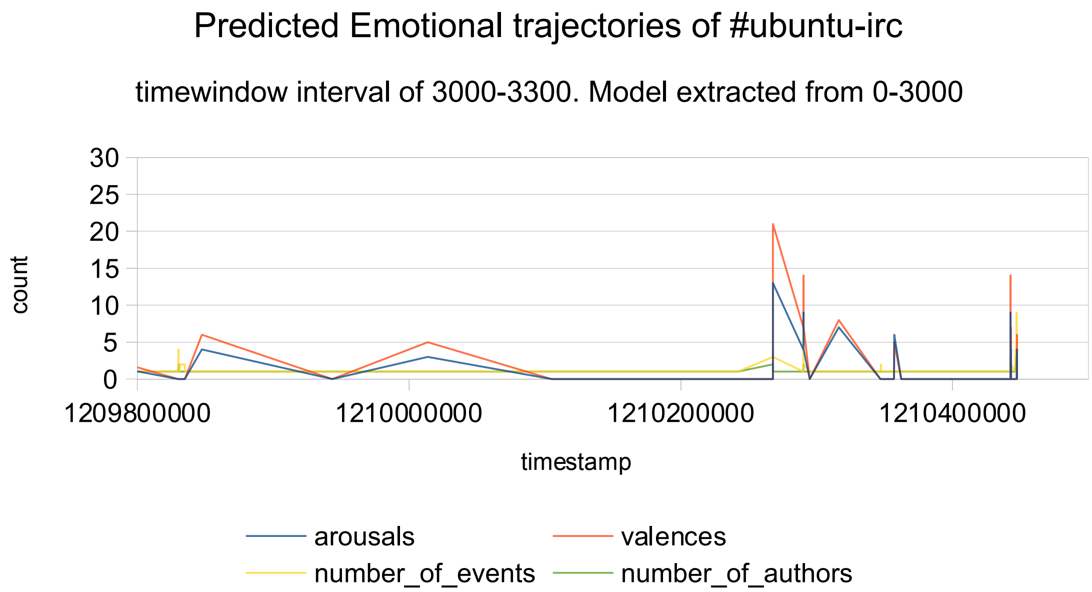

Figure 11. Predicted Emotional trajectories of #irc-ubuntu on timewindow interval of 3000-3000

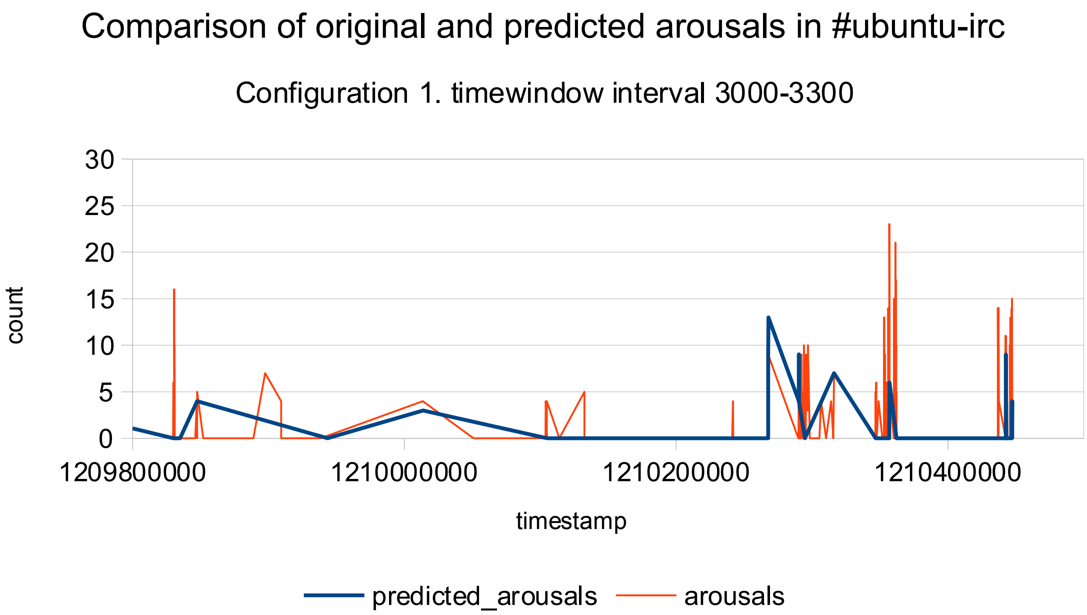

Figure 12. Comparison of original and predicted arousals in #ubuntu-irc. Timeinterval 3000-3000. Configuration 1

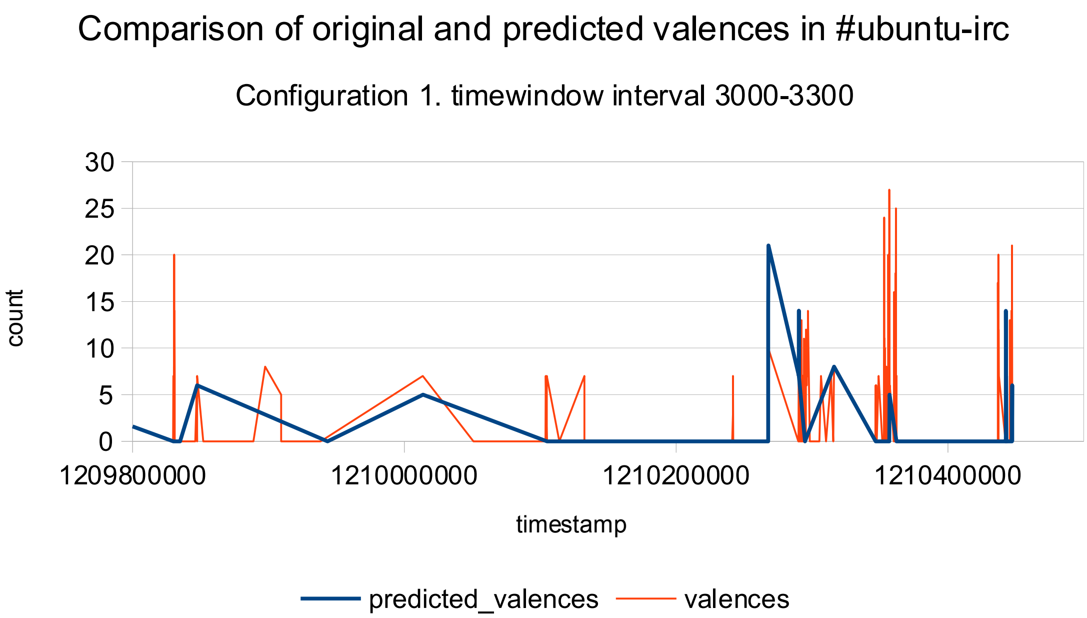

Figure 13. Comparison of original and predicted valences in #ubuntu-irc. Timewindow interval 3000-3300. Configuration 1

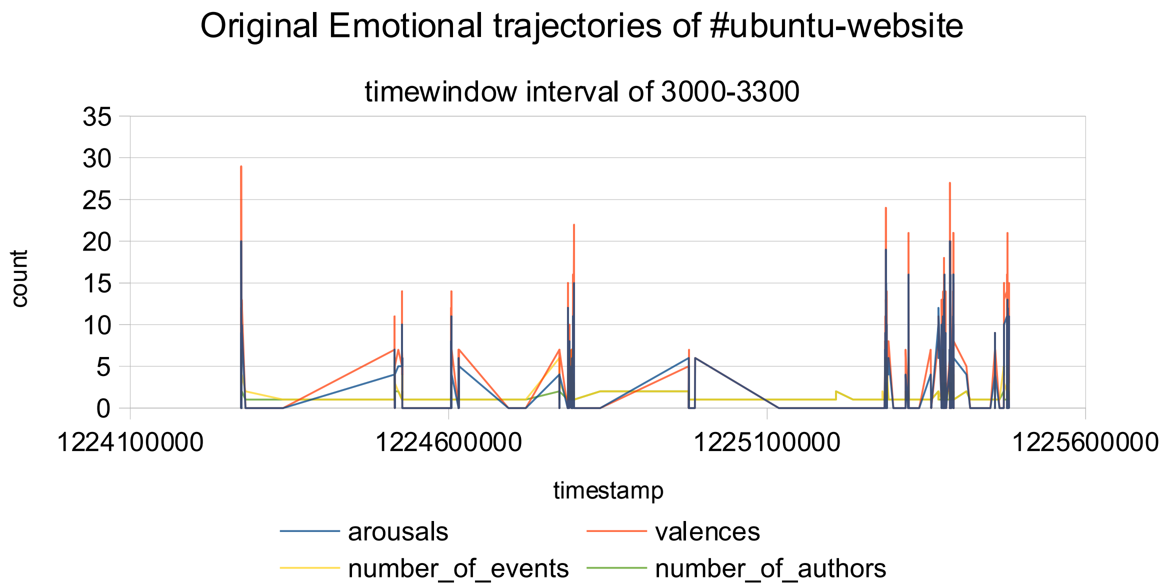

Figure 14. Original Emotional trajectories of #irc-website on timewindow interval of 3000-3000

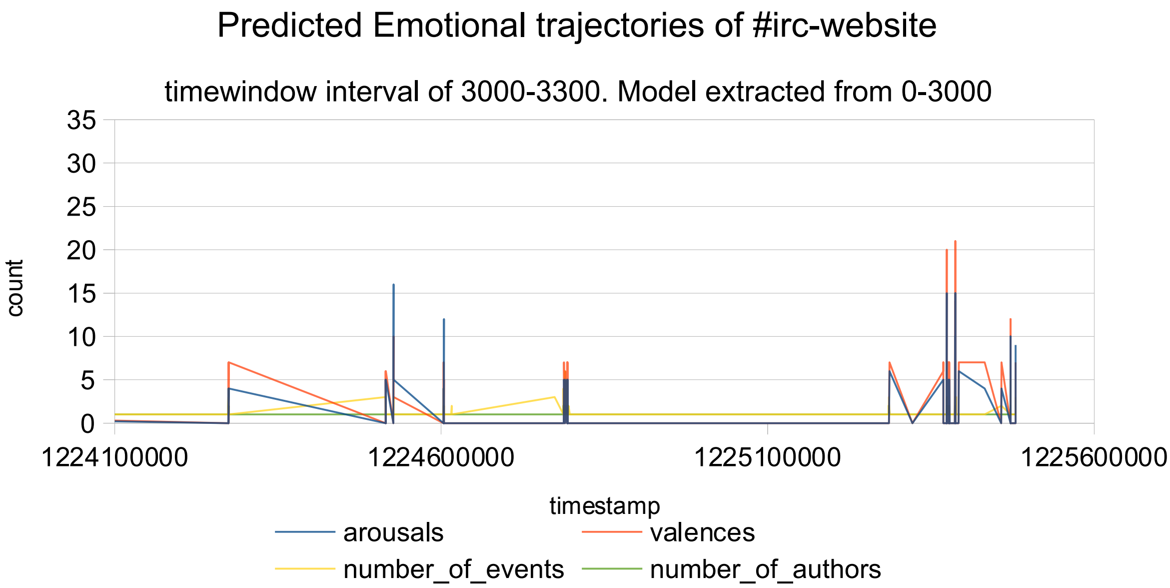

Figure 15. Predicted Emotional trajectories of #irc-ubuntu on timewindow interval of 3000-3300. Derived from 0-3000. Configuration 1

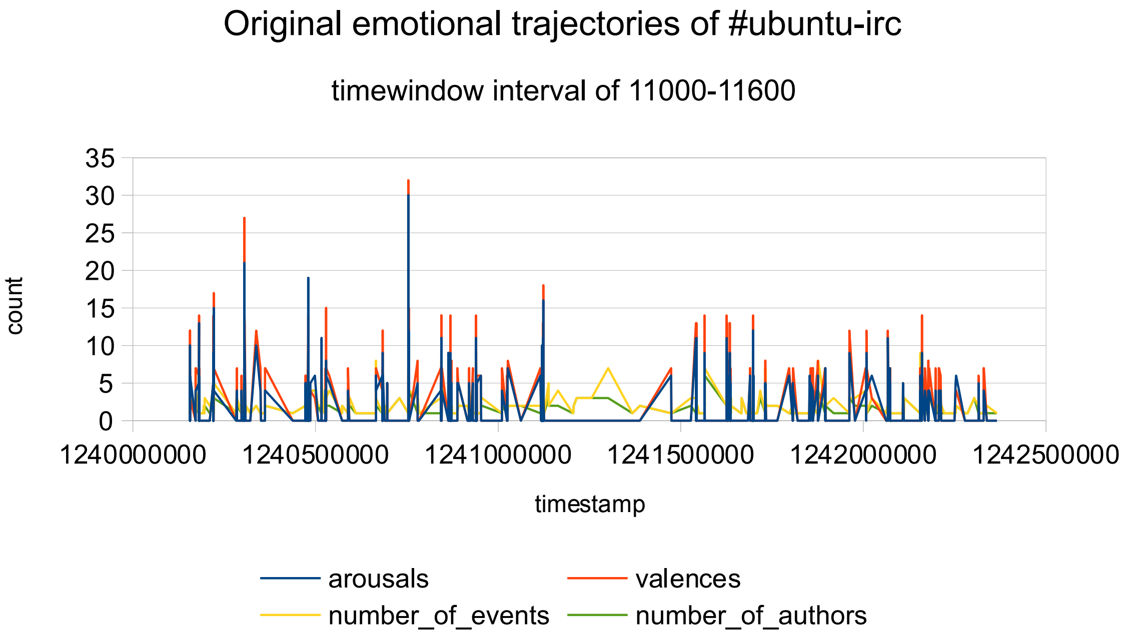

Figure 16. Original Emotional trajectories of #irc-ubuntu on timewindow interval of 11000-11600.

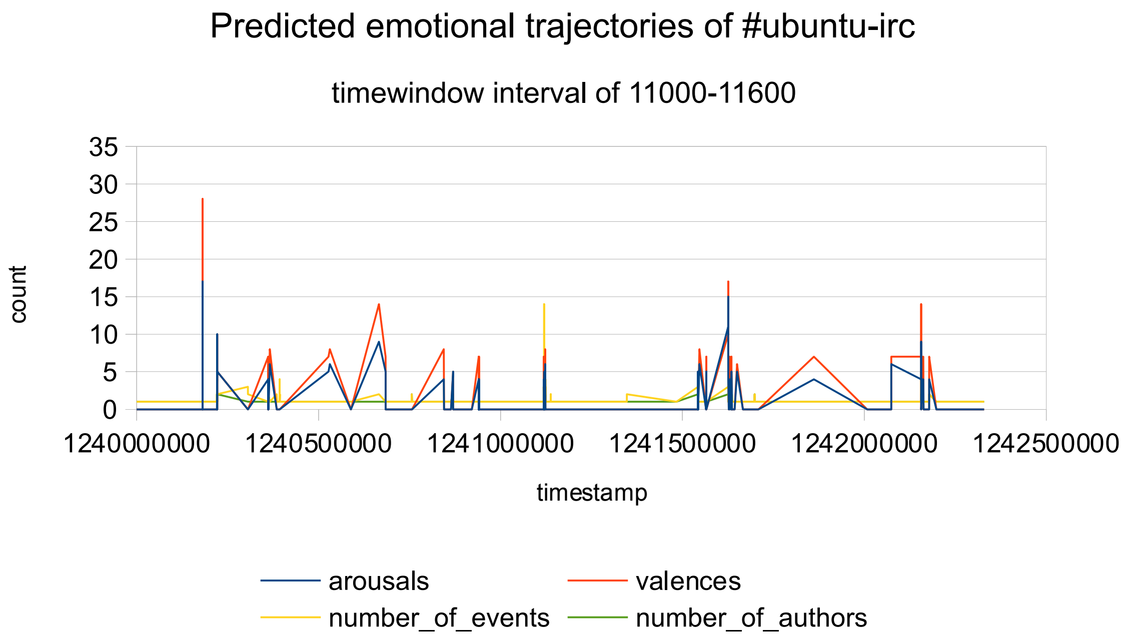

Figure 17. Predicted Emotional trajectories of #ubuntu-irc on timewindow interval of 11000-11600. Derived from 5000-6000. Configuration 1

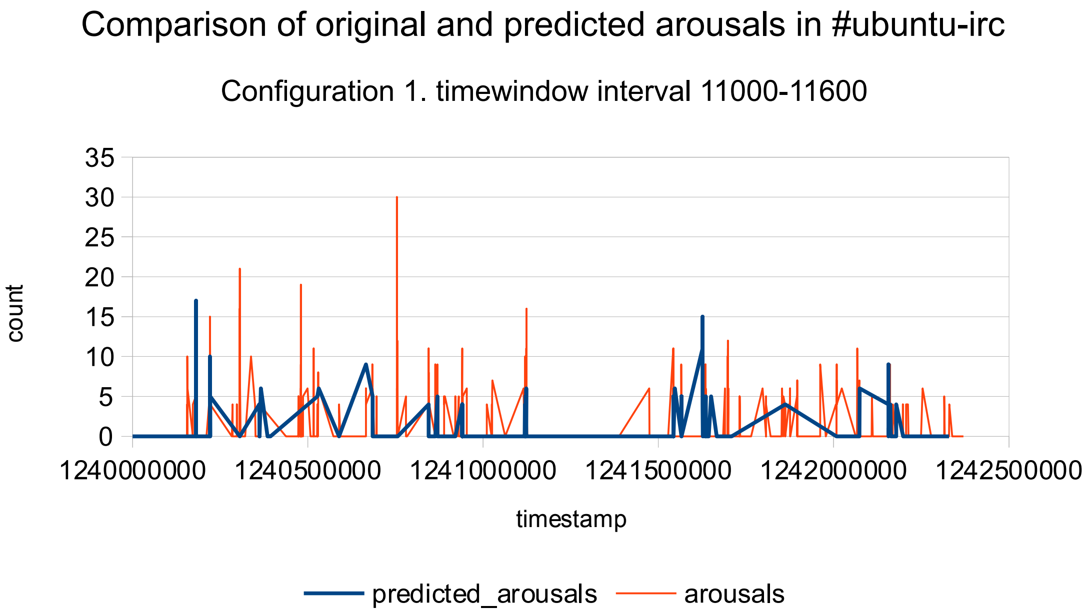

Figure 18. Comparison of original and predicted arousals in #ubuntu-irc. Timewindow interval 11000-11600. Configuration 1

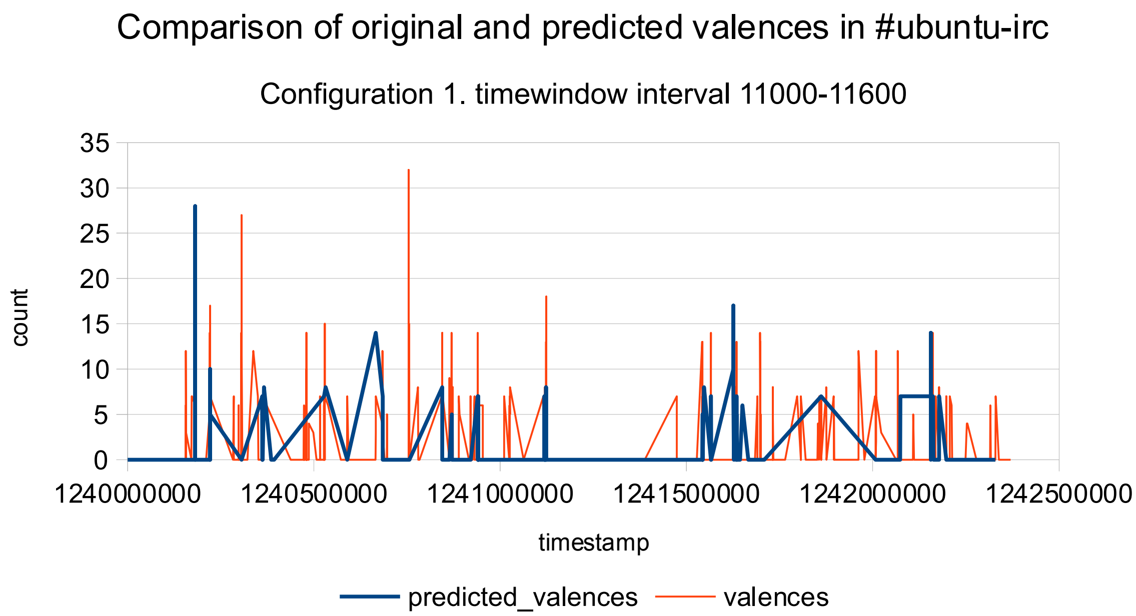

Figure 19.Comparison of original and predicted valences in #ubuntu-irc. Timewindow interval 11000-11600. Configuration 1

As you can see on the graph measurements above, we compared original emotional trajectories with the predicted ones.

The first measurement was done on #irc-ubuntu channel, where the model was derived from the timewindows interval of 0-3000. Simulation was running on the timewindows 3000-3300. Looking at both: Figure 10 with Figure 11 you can see that the positions of the peaks are very similar, as well as the *predicted* emotional trajectories (Figure 11) slightly copies the *orignal* emotional trajectories (Figure 10).

The same applies for other measurements shown in Figure 14 and Figure 15, where we derived the model from the same time window interval (0-3000) and run the simulation on 3000-3300.

Simulations of Configuration 2

Events	#ubuntu-irc	#ubuntu-website	#ubuntu-laptop
Total:	809	689	1384
Correctly predicted:	62	132	1
Incorrectly predicted:	90	555	24
Not predicted:	747	81	1383
Events with recipients:	139	75	15

Table 8. Simulation stats for IRC channels. Model extracted from 0-3000. Simulations run on 3000-3300. ***Configuration 2***

In the Table 8, you can see that the performance of the model evaluation improved in range of 20-30% by the change of configuration data. The number of correct predictions in #ubuntu-irc increase by almost 20% and the number of correct predictions in #ubuntu-website increased by almost 30%.

However, one should be careful of decreasing the strictness of the model, as it could lead into non-valuable data of predictions. Increasing the *emotional epsilon* setting means that after some threshold the model will no longer check, whether an event occurred with a specific emotional value, but it will only check whether an event happened or not. In that case we can not longer talk about emotional predictions, but "reaction" predictions instead.

The following figures (Figure 20, 21, 22, 23) behave very similarly to the behavior we just described above, so the predicted emotional trajectories are consistent with the real one. Those figures can be described analogical to the previous ones.

Interesting behavior is in the Figure 24, 25 where the comparison of #ubuntu-laptop simulation can be seen. Figure 24 shows, the already mentioned data characteristic about #ubuntu-laptop channel, where a lot of new users appears and do not talk too much (especially emotively). This can be seen in the characteristic of this figure, when the number of authors and the number of events is much more bigger than the total amount of all valences or arousals. Therefore, there is not a single peak in the predicted values of #ubuntu-laptop simulation.

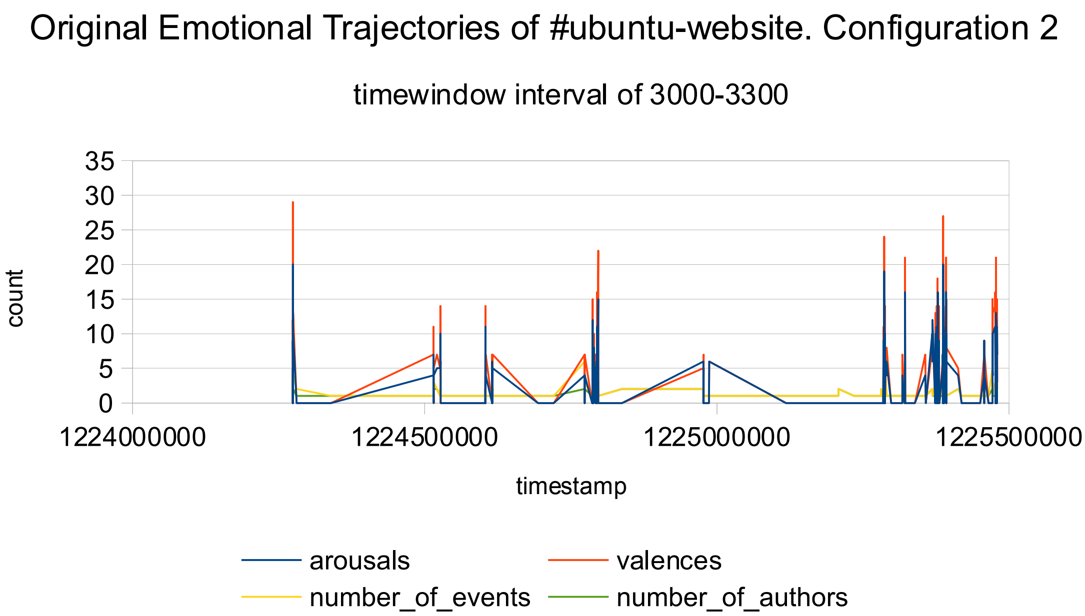

Figure 20. Original Emotional Trajectories of #ubuntu-website. Configuration 2. Timeinterval of 3000-3300

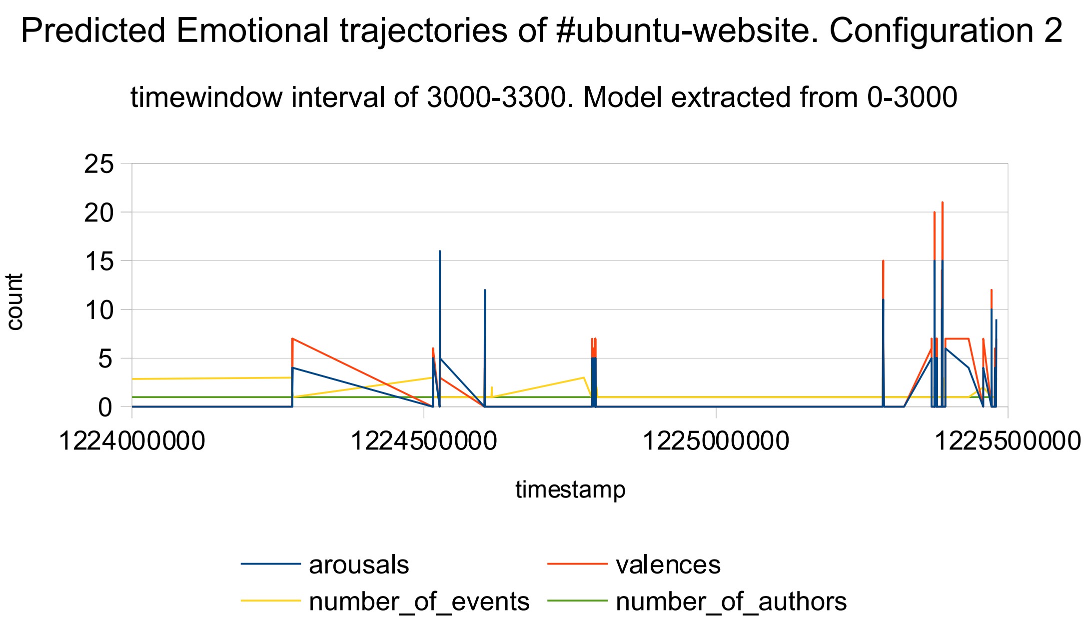

Figure 21. Predicted Emotional trajectories of #ubuntu-website. Configuration 2. Timewindow interval of 3000-3300

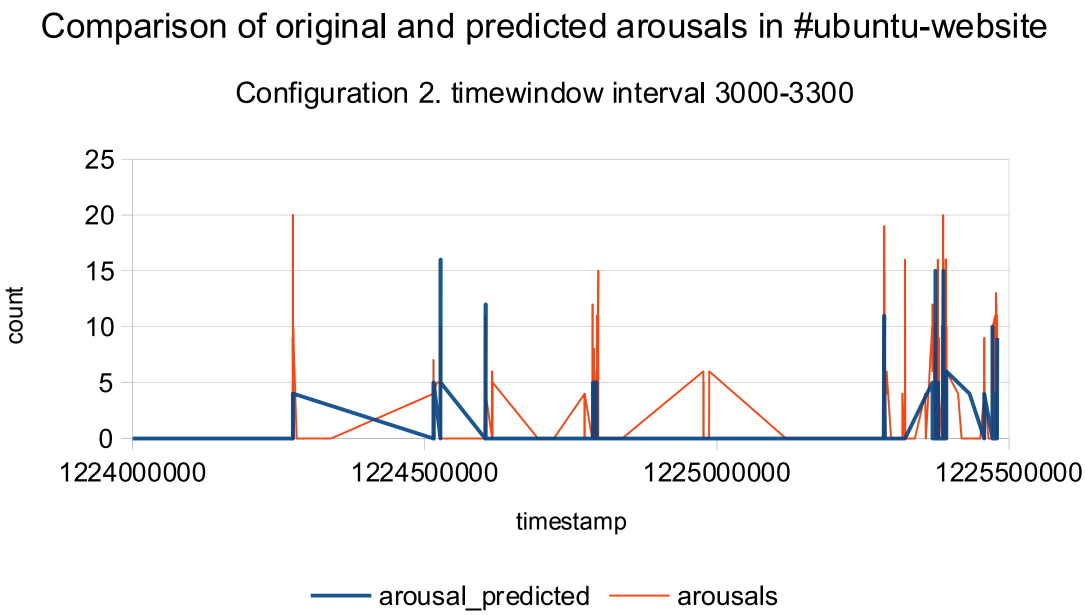

Figure 22: Comparison of original and predicted arousals in #ubuntu-website

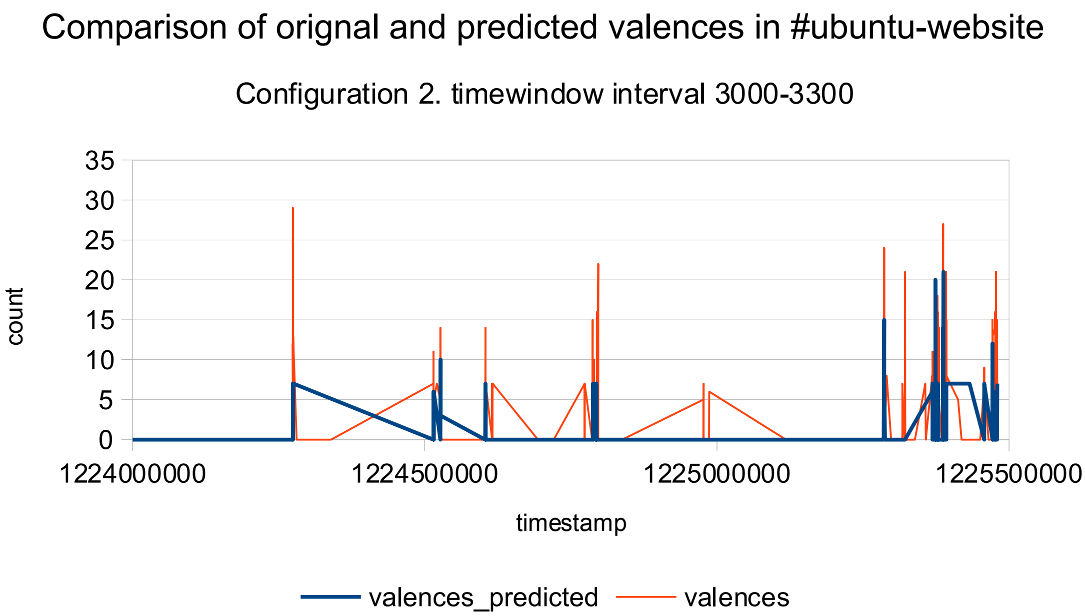

Figure 23. Comparison of original and predicted valences in #ubuntu-website

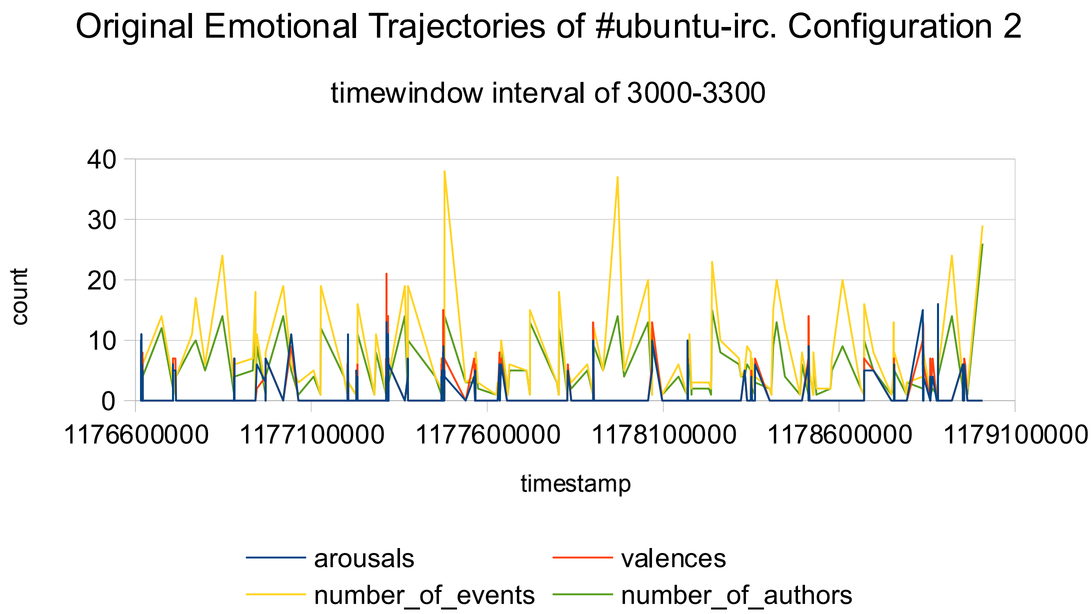

Figure 24. : Original Emotional Trajectories of #ubuntu-irc. Configuration 2. Timewindow interval of 3000-3300

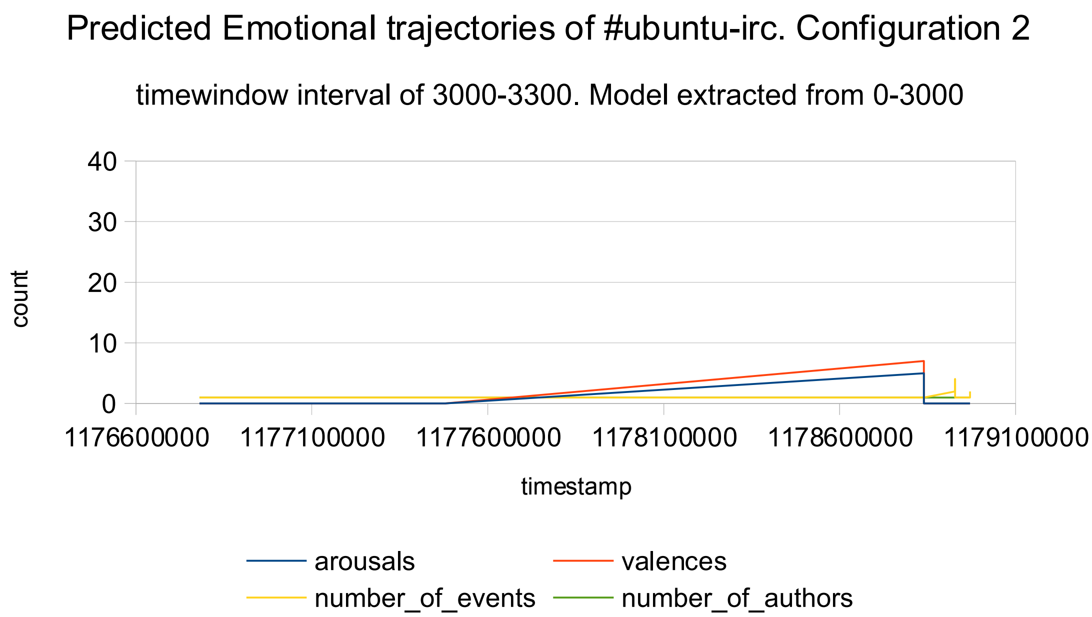

Figure 25.: Predicted Emotional trajectories of #ubuntu-irc. Configuration 2. Timewindow interval 3000-3300.

In this section, we summarized and discussed all the simulations we performed on the annotated datasets of the IRC channels. Except for one dataset (#ubuntu-laptop), the type and the results of the evaluation we performed is consistent with the hypothesis of this work.

4.1 Stochastic theoretical model

One of our hypothesis claims that the MAS model have better performance than a pure stochastic model that is based on probabilities observed on the data. Generally, the performance of a model is measured by counting the number of correctly predicted events that happened in the environment and the trigger is clear.

The event is predicted correctly if the event is similar to a prediction. This means:

- the authors of predicted and real event must be the same
- the difference of valences and arousals between predicted and real event is smaller than Epsilon_Emotional
- time difference (delta) must be smaller than epsilon_time

Applying the “correctness” on the results of a model we can get these numbers:

- number of correctly predicted events
- number of incorrectly predicted events (events for which there is no suitable correct event for a prediction)
- number of not predicted events (events happened but were not predicted)

The first step of the stochastic model for a trigger event processing is considering whether an agent reacts on the event or not. This value we get by calculating the probability of agents replying on trigger events, where a trigger event is a message in an environment which “recipient” is the agent. This probability is called the *reply event* probability and we easily get it by dividing the number of all replies by the number of all trigger events. We define the reply event probability as:

$$P_r = \frac{|reply\ events|}{|trigger\ events|}$$

The probability that an agent does not reply is:

$$P_{nr} = 1 - P_r$$

The reply prediction of the agent is defined by three parameters: time, valence & arousal. The **time** is easily calculated by taking the local time of the trigger event and adding a fixed time constant to it, e.g. (10 minutes). This means that the predicted event is considered as a correct prediction if the real event is going to happen in next *epsilon_time* after the trigger.

The valence & arousal is a different case. The possible values for both are from interval <-5, 5> (ANEW) or <0, 10> (SentiStrength). The values for both ANEW or SentiStrength do not need to be integer values. In the case of valence or arousal, if the value is a real number, we round it into the integer one.

For our purposes we use <0, 10> for both values. (note that ANEW values could be easily transformed to <0, 10> interval by adding 5 to every value)

By taking all valence & arousal together we get a possibility of 121 different values:

- *valence* possible values: 11
- *arousal* possible values: 11
- *valence&arousal*: 11 x 11 = 121

However, the real valence & arousal data distribution in the observing environment is not evenly distributed. By observing the input environment we can see that the probabilities for some values is very close to 0 or is 0 at all, where on the other hand some of the probabilities are significantly greater. For that case, we calculate the distribution of valence & arousal among the whole range of possible values. We define the probability of an event with valence&arousal as:

$$P_{va}ij = \frac{|events\ where: arousal = i \wedge valence = j|}{|reply\ events|}$$

where i, j is from <0, 10> and i, j *in **N***

$P_{va}ij$ is calculated simply by dividing the number of all replies by the number of

$P_{va} ij$ replies.

The sum of all $P_{va} ij$ **,** where i, j is from <0, 10> equals 1.

The probability that a reply event with valence **i** and arousal **j** is predicted correctly is:

$$P_{pred} ij = P_r \times P_r \times P_{va} ij \times P_{va} ij$$

The probability that any reply event is predicted correctly is:

$$P_{pred} = \sum_{i=0}^{10} \sum_{j=0}^{10} P_{pred} ij$$

Application of the stochastic model on the data

For our purposes we take the all history into account and not only the learning part as we do no need further data for validation. Generally, a model based on the learning which already includes testing values has a better performance than a model based on the learning which excludes the testing vales but carries with it the danger of over-fitting.

We apply the stochastic model on the data of the IRC Ubuntu called #ubuntu-irc in the history since 2008-01-16 to 2010-07-18. (IRC-2200Days.#ubuntu-irc.2008-01-16.2010-07-18.h5)

The number of all trigger events in the dataset is 45908. The number of all reply events is 10628. Hence:

$$P_r = \frac{10628}{45908} = 0{,}2315$$

In the following table (Table 9) we can see the distribution of event replies according to the valence and arousal of an event. The total number of all replies is 10628. From the table we can see that most events 7628 were 0-values. However, we can can see a kind of grouping all emotional values around the middle of the spectrum, which means arousal-4, valence-4.

Arousal/ valence	0	1	2	3	4	5	6	7	8	9
0	7624	0	0	0	0	0	0	0	0	0
1	0	0	0	0	2	5	7	2	0	0
2	0	0	2	0	15	30	51	3	1	0
3	0	0	0	9	17	37	55	5	0	0
4	0	0	0	10	58	37	13	0	0	0
5	0	0	11	107	274	146	10	1	0	0
6	0	0	0	7	931	255	9	0	0	0
7	0	0	6	12	270	458	55	6	0	0
8	0	0	0	0	0	30	45	12	0	0
9	0	0	0	0	0	0	0	0	0	0

Table 9. The emotional distribution of event replies according to arousal and valence on IRC- 2200Days.#ubuntu-irc.2008-01-16.2010-07-18.h5 dataset

In the next table (Table 10), we see the probability distribution of emotional event replies. The probability is counted by dividing the number of events with specific valence&arousal (e.g: 7624, 0, 15, 40,..) by the number of all replied events (i.e. 10628)

Arousal/ Valence	0	1	2	3	4	5	6	7	8	9
0	0.7174	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0000	0.0000	0.0000	0.0000	0.0002	0.0005	0.0007	0.0002	0.0000	0.0000
2	0.0000	0.0000	0.0002	0.0000	0.0014	0.0028	0.0048	0.0003	0.0001	0.0000
3	0.0000	0.0000	0.0000	0.0008	0.0016	0.0035	0.0052	0.0005	0.0000	0.0000
4	0.0000	0.0000	0.0000	0.0009	0.0055	0.0035	0.0012	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0010	0.0101	0.0258	0.0137	0.0009	0.0001	0.0000	0.0000
6	0.0000	0.0000	0.0000	0.0007	0.0876	0.0240	0.0008	0.0000	0.0000	0.0000
7	0.0000	0.0000	0.0006	0.0011	0.0254	0.0431	0.0052	0.0006	0.0000	0.0000
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0028	0.0042	0.0011	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 10. Probability of all reply events ($P_{va} ij$) according to specific arousal&valence on IRC-2200Days.#ubuntu-irc.2008-01-16.2010-07-18.h5 dataset..

Hence, with knowing the exact values of P_r and $P_{va} ij$, we calculate P_{pred} :

$$P_{pred} = 0.028217 \quad ,$$

So the probability of the stochastic model to correctly predict a reply event on IRC-2200Days.#ubuntu-irc.2008-01-16.2010-07-18.h5 dataset is $P_{pred} = 0.028217$.

4.1.1 Comparison with the MAS model

For the evaluation of the MAS model we used IRC-2200Days.#ubuntu-irc.2008-01-16.2010-07-18.h5 dataset. We used learning data from the time 2008-01-16 until 2008-03-13 (almost two months [2000 timewindows]).

As a model configuration we used the *epsilon_emotional* = 0. This means that the predicted event had to have the same valence & arousal as a real event. And we can say that under these settings the MAS model is very strict in evaluation of correctness of predicted events and the most error prone.

As a testing data we used the interval from the end of the day 2008-03-13 to 2008-05-03. The data we get from the model is:

number of all events: 3222

number of events with known trigger (reply events): 573

number of correctly predicted events: 204

number of incorrectly predicted events: 676

According to the stochastic probabilities that we proposed above, the correctness of the stochastic model will be:

number of correctly predicted events:

$$|reply\ events| \times P_{pred} = 573 \times 0.028217 = 16.17$$

Here you can see, that the correct events prediction by the MAS (204 correct prediction) is much more effective as the stochastic model does (16 correct predictions). This proves the hypothesis that the evaluation performance of correct predictions by the MAS model is better than the stochastic model. However we can see that the number of incorrect predictions is greater in the MAS model. For our purposes the number of correct predictions is more important than the number of incorrect predictions (from the perspective of bot-support and collective emotions).

Here a “collective emotion” evaluation is needed for better measuring of a “mass” predictions.

4.2 Discussion

The quantity of predicted events (no matter whether correct or incorrect) is related to the quantity of sender-recipient information in the data characteristic. One of the primary inputs that the model needs to predict a user's behavior is to detect a trigger (message) which is aimed to the first user. With the absence of this information, the model's effectiveness is rapidly decreased even it has its own heuristics how to detect a recipient of a message if the sender is not defined.

The quantity of incorrectly predicted events is also influenced by the “saying goodbye” phenomenon. Users in the end of the chatting use to say “bye” or any other farewell to indicate that they are leaving or that the discussion is finished. Here is a good example:

<John> That was a dough to solve this problem.

<Jane> Yes... thanks

<Jane> John: Thank you very much for your help and have a nice sunny rest of the day.

<John> Jane: have a nice day too

... there was not a message from Jane next 10 minutes …

Text 4. IRC channel, "saying-goodbye" example

In Text 4, you can see that John's message to Jane has an emotional value and our model automatically expects that it will elicit an action in Jane. However it should check whether the John's last message is a "bye message" and Jane was not going to respond on it. Such situations could be one of the reasons when the model predicts events that are not going to happen. Message to Jane has an emotional value and our model automatically expects that it will elicit a reaction in Jane. To reduce the number of incorrectly predicted events the model should take the dialog act class of a message into account.

4.3 Future Ideas

Sender – Recipient detection

If there is a communication in the environment between few users and no one of users explicitly mentioned a name of any other users. The model should be able to distinguish that a user message is related to the all users that are currently communicating.

Newbies attack

According to Chapter 3.3.2, in some situations a lot of newbies in a short period of time come to the channel and ask one or two questions and leave the channel. From such an observation we can deduce that something important happened and new users will come to do the same. Here we could adapt the model to predict joins of new users.

Different environment

We created and proposed the model to be suitable for the IRC chat communication. However, the model could be adapted to work on other, asynchronous online communication (blogs, forums).

Different size of timewindow

In the whole work we used a timewindow defined as a 60 seconds timeframe. Changing the size of this timewindow could lead into slightly different results and also could be used in a simulations running on much bigger datasets and simulation time frames. Increasing the size of the timewindow can lead into less-strict evaluations of time-distance of predicted events.

5 Conclusion

In this work, we dealt with the phenomenon of communication in online communities and the emotions which are shared by the users of such communities. We started with an overview of possible suitable agent models and emotional theories on which a computational model of emotional interaction online could be based. We decided to refer to appraisal theories of emotion and, in particular, to one computational model based on an appraisal theory, called WILL. By formally defining the agents and their emotional behavior, as part of the Multi-agent System we used, we developed a framework and the mechanisms suitable for one approach to modeling collective emotions and emotional trajectories of online behavior of single individuals. This model already proves useful as input for a decision mechanism of conversational systems. In this thesis, we evaluate the results of simulation runs of this model in comparison to the reference data it is based on. To further extend its usefulness, we propose using this model to derive a default settings for agents, a 'personality', that represent new users joining an environment based on clustering the agents for already observed users.

According to the results we got, we showed that our hypothesis & goals for creating the model applies: the emotional predictions were consistent with the reference data. Even though we created the model on a simple idea of a trigger-reaction scheme, where the context of communication is not taken into account. We also saw that the performance of the model strongly depends on the characteristics of the modeled community and its users. Online IRC communities with many newbies and few stable long-term users show bad results in the predictions as there is not enough input data from the model for deriving agents' personalities.

Further evaluation showed that the model also fared very well in comparison with a basic stochastic model based on static probabilities derived from the reference data.

Finally, we point out several ways in which the model could be improve in future iterations (e.g. different representation of emotional state). The robustness of the model, however, can rapidly deteriorate with such changes, so that they will need to be weighed against one of the goals we postulated: *"to balance the need for a suitable degree of complexity with the need for a simple and fast simulation"*.

The model and simulation system was implemented as a Multi-agent system, which can be easily run on different datasets with different simulation parameters, ready for further simulation experiments with different kinds of data and online communities.

6 References

Arnold, M.B., (1954). Feelings and Emotions As Dynamic Factors in Personality Integration. In: Arnold, M.B., Gasson, J. (1954) The human person, pp. 294-313. Ronald Press-New York

Bradley, M.M., Lang, P.J., (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Carver, C.S., Scheier, M.F., Weintraub, J.K., (1989). Assessing coping strategies: a theoretically based approach. Journal of Personality and Social Psychology, 56(2):267-283.

Chmiel, A., Holyst, J.A., (2010). Flow of emotional messages in artificial social networks. International Journal of Modern Physics, C 21(5), 593-602.

Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., Holyst, J.A., (2011). Collective emotions online and their influence on community life, ĘPLoS ONE

Chmiel, A., Sienkiewicz, J., Paltoglou, G., Buckley, K., Thelwall, M., & Holyst, J. A., (2011). Negative emotions boost users activity at BBC Forum, Physica A, 390(16), 2936-2944.

Cowie, R., (2007). D3i Emotional life Terminological and conceptual clarifications.Contract, (October), Available at HUMAINE Portal: http://emotion-research.net/projects/humaine/deliverables/D3i%20final.pdf ,pp.1-50.

CyberEmotions. (2012). CyberEmotions project's webpage. 28th February, 2012.

Retrieved from http://www.cyberemotions.eu

Derks, D., Bos, A. E.R., VonĘGrumbkow, J., (2008). Emoticons and online message interpretation. Soc. Sci. Comput. Rev., 26(3):379-388.

Doce, T., Dias, J. Prada, R., Paiva A., (2010). Creating Individual Agents through Personality Traits. in Intelligent Virtual Agents Springer-Verlag Berlin Heidelberg.

Ekman, P., (1999). Basic Emotions. In: T. Dalgleish, T. Power (Eds.) The Handbook of Cognition and Emotion Pp. 45-60. Sussex, U.K.: John Wiley & Sons, Ltd.

Ellsworth, P.C., Scherer, K., (2003). Appraisal processes in emotion. In: Davidson R.J. et al. (Eds.), Handbook of Affective Sciences, pp. 572-595. Oxford New York: Oxford University Press.

Folkman, S., Lazarus, R. S., (1980). An analysis of coping in middle-aged community sample. Journal Health Social Behavior, 21, 219-239.

Frijda, N.H., (1986). The Emotions., Cambridge University Press. ISBN: 9780521316002

Frijda, N.H., Swagerman, J., (1987). Can computers feel? Theory and design of an emotional system. Cognition and Emotion., I(3) pp.235-257

Frijda, N.H., (2007). The Laws of Emotion. Lawrence Erlbaum Associates, Publishers, Mahwah NJ USA, London UK EU.

Garas, A., Garcia, D., Skowron, M., Schweitzer, F., (2012). Emotional persistence in online chatting communities. Scientific Reports, 2.

Gilbert, G.N., Conte, R., (1995). Artificial societies: the computer simulation of social life. London: UCL Press.

Gratch, J., Marsella, S. (2001). Tears and fears: modeling emotions and emotional behaviors in synthetic agents. In Agents, pp. 278-285.

Gratch, J., Marsella, S. (2004). A domain-independent framework for modeling emotion. Cognitive Systems Research, 5(4):269-306.

Hatzivassiloglou, V., McKeown, K.R., (1997). Predicting the Semantic Orientation of Adjectives. Proceedings of the Joint ACL/EACL Conference, pp. 174-181

Hou, W., Jin, Y., Yao, F., (2008). Research on Multi-Agent System (MAS) Model Based on Flow of Affective Computing (FAC). Pervasive Computing and Applications 2008 ICPCA 2008 Third International Conference on (Vol. 2, pp. 869-874).

Huhns, M.N., Stephens, L.M., (1999). Multiagent Systems and Societies of Agents. In: G. Weiss (Ed.), Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, pp. 79-120., The MIT Press, Cambridge, MA, USA.

Kensinger, E.A., (2004). Remembering emotional experiences: the contribution of valence and arousal., Rev. Neurosci., 15(4):241-51.

Kim, S., Hovy, E., (2004). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (COLING '04). Association for Computational Linguistics, Morristown, NJ, USA

Liu, B., (2010). Sentiment analysis and subjectivity. In: Indurkya, N., Damerau, F. J.(Eds.), Handbook of Natural Language Processing, Second Edition.

Marsella, S., (2003). Modeling coping behavior in virtual humans: Don't worry, be happy. In In AAMAS 2003, pages 313-320

Marsella, S., Gratch, J., Petta, P., (2005). Computational Models of Emotion. In: Scherer, K.R., Bänziger, T., Roesch, E. (Eds.) A blueprint for an affectively competent agent:.

Oxford: Oxford Press

Mascarenhas S., (2010). A Dimensional Model for Cultural Behavior in Virtual Agents, pp.552-574. In Joao Dias, Rui Prada, Ana Paiva, . Taylor, Francis, Inc.

McRorie, M. Sneddon, I. de Sevin, E. Bevacqua, E. and Pelachaud, C. A., (2009). Model of Personality and Emotional Traits. in Intelligent Virtual Agents. Berlin Heidelberg: Springer-Verlag

Mehrabian, A., (1978). Measures of individual differences in temperament. Educational and Psychological Measurement, 38, 1105-1117.

Mehrabian, A., (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Current Psychology: Developmental, Learning, Personality, Social, 14, 261-292.

Miller, G.A., (2010). 27th October. About WordNet. The Trustees of Princeton University. Retrieved from http://wordnet.princeton.edu/

Moffat D., Frijda, N.H., Phaf R.H., (1993). Analysis of a model of emotions. In: Sloman A. et al. (Eds.) Prospects for Artificial Intelligence., pp.219-228 16, IOS Press, Amsterdam;

Moffat D., Frijda, N.H., (1995). Where there's a Will there's an Agent. In: Wooldridge M., Jennings

Moffat, B., (1997). Personality Parameters and Programs, pp. 120-165, Number 1195 in Lecture Notes in Artificial Intelligence. Springer-Verlag

Moss, S., Davidsson, P., (2000). Multi-Agent-Based Simulation, Second International Workshop, MABS 2000, Boston, MA, USA, July, 2000, Revised and Additional Papers. Volume 1979 of Lecture Notes in Computer Science., Springer-Verlag (2001)

Nwana, H.S., (1996). Software Agents: An Overview Knowledge Engineering Review, 11(2):205-244.

Parrott, W., (2001). Emotions in Social Psychology, Philadelphia: Psychology Press,

Picard, R., (1997). Affective Computing. In: Stork, D. (Ed.), United States: The MIT Press.

Rank, S., (2010). Docking agent-based simulation of collective emotion to equation-based models and interactive agents. In Proceedings of the 2010 Spring Simulation Multiconference, New York, NY, USA. ACM.

Rawlinson, L., Hunt, N., (2009). CNN, Jackson dies, almost takes Internet with him, June 26th 2009. retrieved from: http://articles.cnn.com/2009-06-26/tech/michael.jackson.internet_1_google-trends-search-results-michael-jackson?_s=PM:TECH

Reilly, S. N., (2006). Modeling What Happens Between Emotional Antecedents and hEmotional Consequents. In Agent Construction and Emotions (ACE 2006). Retrieved from: http://www.ofai.at/~paolo.petta/conf/ace2006/

Reisenzein, R. (2001). Appraisal processes conceptualized from a Schema-Theoretic perspective: Contributions to a process analysis of emotions. In: Scherer, K.R., Schorr, A., and Johnstone, T., (Eds.), Appraisal Processes in Emotion, chapter 10, pp. 187-201. Oxford University Press.

Revelle, W., (1995). Personality Processes, Annual Review of Psychology, 46:295-328.

Revelle, W., Scherer, K.R., (2010). Personality and emotion. In the Oxford Companion to the Affective Sciences, Oxford University Press

Roseman, I.J., Smith, C.A., (2001) Appraisal Theory: Overview, Assumptions, Varieties, Controversies. In: Scherer, K.R., Schorr, A., Johnstone, T. (Eds.) Appraisal Pro- cesses in Emotion: Theory, Methods, Research, pp. 3-19. Oxford New York: Oxford University Press

Rousseau, D., Hayes-Roth, B., (1998). A Social-Psychological Model for Synthetic Actors. in second international conference on Autonomous agents, ACM New York

Scherer, K. R., (2001). Appraisal Considered as a Process of Multilevel Sequential Checking. In: K. R. Scherer, A. Schorr, T. Johnstone (Eds.), Appraisal Processes in Emotion: Theory, Methods, Research., pp. 92-120, Oxford University Press.

Scherer, K.R., Wranik, T., Tran, V., Scherer, U., (2004). Emotions in everyday life Social Science Information 43, No. 4, 499-570

Scherer, K.R., (2005). What are emotions? And how can they be measured?, Social Science Information 44(4):695-729. Scherer, K., Goldsmith, H. (Eds.), Handbook of Affective Sciences., pp. 572-595, New York: Oxford University Press,

Schlosberg, H., (1954). Three Dimensions of Emotion, Psychological Review 61: 81-8.

Schwarz, N., Bohner, G., (2001). The construction of attitudes. In Tesser, A. and Schwarz, N., editors, Blackwell Handbook of Social Psychology: Intraindividual Processes, pp. 436-457. Blackwell, Malden, MA.

Schweitzer, F., (2007). Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences. Springer Publishing Company, Incorporated.

Schweitzer, F., Garcia, D., (2010). An agent-based model of collective emotions in online communities. The European Physical Journal B - Condensed Matter and Complex Systems, 77:533-545.

Smith, C. A., Kirby, L., (2000). Consequences require antecedents: Toward a process model of emotion elicitation. In: J. P. Forgas (Ed.), Feeling and Thinking: The role of affect in social cognition, pp. 83-106, Cambridge University Press.

Smith, C.A., Lazarus, R., (1990). Emotion and Adaptation. In: L. A. Pervin (Ed.), Handbook of Personality: theory & research, pp. 609-637, NY: Guilford Press.

Swagerman, J., (1987). The Artificial Concern Realization System ACRES: A Computer Model of Emotion. Doctoral dissertation, University of Amsterdam

Tanguy, E., Willis, P., Bryson, J. (2009). Simplifying the Design of Human-Like Behaviour: Emotions as Durative Dynamic State for Action Selection

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A., (2010). Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12), 2544-2558

Thelwall, M., Buckley, K., and Paltoglou, G., (2012). Sentiment strength detection for the social Web. Journal of the American Society for Information Science and Technology, 63(1), 163-173.

Wooldridge, M., (1995). This is MYWORLD: the logic of an agent-oriented testbed for DAI. In Intelligent Agents: Theories, Architectures and Languages, Wooldridge, M., Jennings, N. R., (Eds.), LNAI Volume 890, pp. 160-178, Springer, Rerlin

Wooldridge, M. (2002). An Introduction to MultiAgent Systems. John Wiley & Sons, 1st edition. Wiley, Chichester, UK. ISBN 0 47149691X.

Wooldridge, M. (2009). An Introduction to Multiagent Systems. Wiley, Chichester, UK, 2. edition. Wiley, Chichester, UK. ISBN 978-0-470-51964-2

Wundt, W. (1880). Grundzüge der physiologischen Psychologie. Thömmes.

Appendix

This thesis comes with an additional software and data distribution. In the package (project.zip), you can find two directories: results and cybabmod.

The *Results* directory contains all the extracted results we got from the simulations and related graphs for them.

Cybabmod is the whole Python code for the implementation of the MAS we proposed in the paper. All the files are documented inline and all important methods of the code are covered by Unit Tests.

For instructions on running the simulations or data extraction, refer to the README file included in the *Cybabmod* directory.

Note that the HDF5 reference datasets that were used in this work cannot be freely distributed as they have been collected as part of the CyberEmotions project, and are therefore not attached to this thesis. The datasets are, however, available and based on freely available data: please feel free to contact the author of the thesis if you wish to get access to the reference data.