Chapter 14
Zooming in: Studying Collective Emotions with Interactive Affective Systems

Marcin Skowron, Stefan Rank, David Garcia, and Janusz A. Holyst

14.1 Introduction

Computer-mediated communication between humans is at the center of the formation of collective emotions on the Internet. In parallel, the influence of artificial systems and services on the pace of information propagation and the selection of topics is growing and this needs to be accounted for in a holistic analysis of the emergence of collective emotions on networks. Interactive artificial systems can play an important role in studying the influences on the formation of collective emotions in cyberspace.

Numerous online services rely on computer algorithms for selecting and prioritizing content displayed to users. This enables websites that collect millions of units of information, constantly updated and served to millions of users in a timely and often personalized fashion. The scope of potential applications of such systems and their effects on personal and collective states can easily be underestimated by the average Internet user. Artificial systems can also actively search and establish contact with users and provide content. This ranges from the
occasional dispatch of email or twitter messages to synchronous natural-language-based chats or task-specific dialogues in a written or oral form. The artificial systems have different functions and identities: from operators and supportive bots on internet communication channels, personal assistants on mobile devices, over ‘non-playable characters’ (NPCs) in online games and ‘chat bots’ applied as means of entertainment, to a range of service and information providers in commercial, health-care and educational contexts. The multi-layer impact of systems with different characteristics of their interactive behavior on the spread of information and emotions, on the formation of social bonds between users as well as on decisions concerning personal or group choices clearly reach beyond the online world.

In this chapter we focus on a subset of artificial systems which can be characterized by their strongly interactive nature: specifically involving direct, natural language based communication with users, paired with the ability to perceive, model, and elicit emotions. We elaborate here that such interactive affective systems (1) play an important role in the scientific exploration of the factors which influence the formation of individual and collective emotions in online setups, (2) show great potential for practical applications and (3) necessitate a discussion on ethical aspects related to their deployment in real world e-communities.

The rest of this chapter is structured as follows: First, we present the motivation for interactive studies of collective emotions and introduce practical examples of application scenarios for artificial interactive systems endowed with an awareness of affective phenomena. We also describe how such systems can be applied for interactive studies of processes related to the formation of collective emotions and provide an overview of the main components related to the computational awareness of emotions in artificial systems. The following section presents the experimental results obtained from a series of experiments with a specific realization of interactive affective systems. Next we enumerate challenges related to application of IASs in multiple users environment and present an overview of a theoretical framework and specific agent-based modelling approach which support the systems’ decision making and influence their interactive and affective characteristics in such environments. Finally, we discuss ethical implications of the application of interactive affective systems for supporting online e-communities.

14.2 An Interactive Approach to Collective Emotions

Recent studies of collective emotions in online communication are focused on post-hoc analysis or modelling of valence exchanged in textual messages and the related interaction patterns in e-communities (Mitrović and Tadić 2011; Chmiel et al. 2011a; Thelwall et al. 2011, 2013; Hillmann and Trier 2012; Garas et al. 2012), as well as physiological responses related to the perception and generation of emotionally charged online content (Kappas et al. 2010). The interactive study of hoxo-communities complements these lines of work by employing affective interactive systems to provide experimental setups in order to (a) support studies
on how interactions with artificial and human agents influence us in online virtual settings (Skowron 2010); (b) gather supplementary data on specified topics from selected groups, enabling the acquisition of relevant background information that is not available to post-hoc analysis (Skowron et al. 2011c); (c) evaluate theoretical models based on quantitative analyses of large data-sets by experimentally reproducing posited effects online (Rank et al. 2013). The approach also enables direct querying of users to discover a range of variables that are likely to influence affective reception, thus extending the scope of analysis to e.g. motivations and stances of particular users.

14.2.1 Examples of Application Scenarios for Interactive Affective Systems

Interactive affective systems (IAS) aim to model the affective dimensions of multi-party interactions, to simulate potential future changes in group dynamics, and, in part based on that information, to suitably respond to utterances both on the content and the affect level (Skowron and Rank 2014). In the following, we outline two general examples of realised application scenarios of IAS to illustrate the range of foreseen activities for such systems. The Affective Interaction Analyser (AIA) focuses on the analysis of interaction patterns, especially in multi-user environments. It is only concerned with the affective content of the exchanged textual messages and the tracking of group-level characteristics and is inactive in terms of interactions with casual users, except for infrequent messages provided to selected users such as e.g. Internet Relay Chat (IRC) channel operators. The Affective Supporter and Content Contributor (ASCC) participates to a moderate extent in ongoing discussions by providing new content related to the discussed topic or the results of affective group dynamic analysis and real-time simulations, communicating with the whole group.

14.2.2 Role of the Method

The role of the proposed method, i.e., interactive studies of e-communities, rests on the interplay between quantitative and qualitative research. This mutually beneficial relation is similar to practice in psychology or sociology: questionnaires and field studies complement statistical data and enable to (a) zoom in to particular groups of subjects, aspects of interactions, or topics of interest; and (b) investigate various hypotheses in direct interactions with users. Current post-hoc methods are limited to the acquisition of large data-sets of online communication paired with the application of algorithms for detecting sentiment expression. The non-interactive approach to studying collective emotions, by its very nature, can only
acquire a limited amount of information, bounded by the state of the art in sentiment classification, information extraction, natural language processing and data mining. Additionally, information posted on the Internet often presents only a distilled summary of users’ opinions on a given topic of interest, frequently missing any explicit statement of affect due to social convention or shared background knowledge, specific to a given community.

The validity of applying artificial IAS in experimental online setups in order to overcome those limitations is grounded in the influence that emotions exert on communication processes even in virtual setups. It has been argued that people treat interactions with certain media (including virtual agents) in the same way as interactions with real humans (Reeves and Nass 1996). Further, for example Forgas (2011) demonstrated that mood influences the disclosure of personal information both in real and virtual setups. Such disclosure is an essential part of human relationship formation (Altman and Taylor 1973). These studies provide the foundation for transferring experimental results obtained in human-computer interaction setups to human-human interaction. Other aspects relevant to the practical implication of experimental results, presented in this chapter and privy to this particular method, relate to the experimental evidence on people’s increased willingness to cooperate with agents that are more human-like [mechanisms similar to kin selection (Hamilton 1964)], and are more engaged in interactions with agents that express emotions (Bates 1994; Lester et al. 1997).

A prerequisite for IAS usable for this method and that can further potentially affect collective states are mechanisms to assess the impact of collective emotions on online and offline processes, described in the following section.

14.3 Computational Awareness of Collective Emotions

Several aspects of the role of emotions and collective emotions in offline communities have been identified, such as: an increased potential for responding to opportunities or threats (Kemper 1991; Preston and Waal 2002); the benefit of emotional similarity for perceiving others’ intentions and motivations (Hatfield et al. 1994; Levenson and Ruef 1992); or validating one’s feelings and appraisals (Locke and Horowitz 1990; Rosenblatt and Greenberg 1991). The implementation of an awareness of the affective dimension of online interactions has significant benefits for the development of artificial systems. The ability to recognise these complex dimensions of human behaviour and related interaction patterns allows for an increased adaptability to social, real-world interactive settings. An understanding of the fundamental aspects of human emotions, caring about specific states of the world, and the subjective assessment of the relevance of any changes in that regard (Ellsworth and Scherer 2003; Petta 2003; Frijda 2007), is a step towards achieving beneficial human qualities such as cooperation, empathy, fairness, or reciprocity that rely on the concern of the well-being of others in artificial systems (Marsella et al. 2010).
In principle, interactive affective systems can communicate with users directly, provide new content to a group of users, and provide reports on group activities identifying interaction patterns. While the interaction is, in this sense, unrestricted, the chosen domain for an IAS is constrained in so far as the systems are concerned mainly with the emotional states of individual users as well as the dynamics of group and collective emotions and employ suitable strategies to keep the conversation going. This restriction is based on the strong influence of emotion in group dynamics, as mentioned above, as well as by consideration of practical feasibility of system complexity. The system, thus, analyses affective aspects of group interactions, utilising information about the network structure, state of collective emotions perceived in an e-community, and the contributions of individuals, and provides reports on the simulation outcomes for the group. In the following, we outline the components that are necessary to achieve this computational awareness of emotion. More detailed description of the components for computational awareness of emotion in IAS and specific methods and tools used for their detection, representation and modelling is presented in Skowron and Rank (2014).

### 14.3.1 Components for Computational Awareness of Emotion in Interactive Systems

A prerequisite for IAS that intentionally study and affect collective states are mechanisms to assess the impact of collective emotions on online and offline processes. There are methods for relating textual expressions to a range of affective dimensions but, additionally, the modelling of large e-communities at different levels that account for the role of emotions has been undertaken. The proposed realisation of computational awareness of collective emotions for IAS is based on these abilities to automatically detect and categorize emotional expressions, to model the dynamics of emotion exchanges in e-communities, and to interact with users directly including affect generation. We distinguish two (overlapping) layers of competencies for IAS: (a) Direct interactions, i.e. dialog analysis in 1-on-1 and multi-user settings, where analysis is focused on the perception of affective dimensions of the dialog and the identification of related aspects such as timing, style, novelty and coherence of contributions; and (b) Social/network interactions, i.e. perceiving, modelling and simulating interactive and affective dynamics in online groups, which includes accounting for the network structure and the roles of individual nodes as well as, for the ASCC scenario, the ability to select peers for interaction. The following list provides an overview of system components relevant for the scenarios outlined above organised along these two layers of competencies.

- **Direct Interactions**
  - Modelling of conversational partners
  - Modelling of self in 1-on-1 interactions
- Affect detection
- Affect generation and affective dialog management
- Management of long-term interactions

• Social and Network Aspects
  - Modelling groups of individuals
  - Modelling of self in online social networks
  - Anticipation of effects of interactions in multi-user environments
  - Modelling and analysis of collective emotions

The development of systems capable to model and potentially influence emotional dynamics in groups based on such an computational awareness of emotion presents new opportunities to support e-communities. Making information on affective dynamics explicit and presenting this information or directly interacting with the group, e.g., by posting relevant novel content, has the potential to increase cooperation in a group or to counteract negative tendencies. In the following, we present the interaction scenarios and the experimental results which were primarily focused on the evaluation of such systems and studies of their impacts on users.

14.4 Interaction Scenarios for Experiments with an Affective Dialog System

The commonality between the experimental setups presented here is that the variants of the Affective Dialog System communicate with users in a predominantly textual modality, rely on integrated affective components for computational awareness of emotion, and use the acquired information to aid generation and selection of responses. The system variants interact with users via a range of communication channels and interfaces that share common characteristics of online chatting. In study 1 this consisted of a three-dimensional virtual reality setting, in studies 2 and 3 a web-chat like interface was used. Detailed description of the experimental systems is presented in the following publications: (Skowron et al. 2011b, 2013, 2014) Here we provide an overview of the experimental setups and main findings from the conducted tests.

14.4.1 Study 1: Virtual Reality WOZ Setting

The focus of the first study was the evaluation of the system in comparison with a Wizard-of-Oz (WOZ) setting as a prerequisite for the following studies. Further, we investigated the practical use of affective cues detected in user input for the generation of affective responses. As setting for the experiment, a Virtual Reality (VR) bar was created: a furnished virtual bar room and a virtual bartender with facial
expressions and a chat interface, see Gobron et al. (2011) regarding the technical aspects of the VR setting.

The experimental setting consisted of the user, represented by an avatar (male or female according to the user’s gender), interacting with a virtual human (male bartender). Each participant interacted four times, 5 min each, randomized order, in 2 × 2 conditions: The conversational partner was either a variant of the system called Affect Bartender (AB) or a Wizard of Oz (WOZ), and the generation of emotional facial expressions was either active or not. In the AB condition, a simulation of thinking and typing speed was introduced to prevent an influence of differences in the response delivery time between the system and the human operator.

### 14.4.2 Study 2: Distinct Affective Profiles

In this study, an artificial affective profile was defined as a coarse-grained simulation of affective characteristics of an individual, corresponding to dominant, observable affective traits, that can be consistently demonstrated by a system during the course of its interactions with users (Skowron et al. 2011c). In this round of experiments, three distinct affective profiles were implemented for the dialog system—labeled as positive, negative and neutral—limiting variations to baseline levels of positive and negative affectivity in personality (Watson and Tellegen 1985). Each affective profile aims at a consistent demonstration of character traits of the system that are described as, respectively:

- polite, cooperative, empathic, supporting, focusing on similarities with a user;
- conflicting, confronting, focusing on differences with a user;
- professional, focused on the job, not responding to expressions of affect.

The study directly addressed research questions about the artificial system’s ability to consistently simulate affective profiles.

### 14.4.3 Study 3: Key Social Processes

Affective profiles simulate human dispositional variations in baseline valence, sometimes referred to as “mood” (Watson and Tellegen 1985) or “affective home

---

1Participants were led to believe that they communicate with a dialog system, while responses are actually provided by a human operator. In the presented experiments, the operator adhered to general guidelines stating the objectives that needed to be achieved during the interactions. These included: providing realistic and coherent responses to the users’ utterances and avoiding utterances that demonstrate an unusual sense of humor or eloquence. Before initiating the experiments, the operator conducted several rounds of pre-test interactions which helped to test and assure a consistency of his communication patterns, in line with the general instructions.
base” (Kuppens et al. 2010) that impact behavior patterns in different communication scenarios and are an important part of implementing different personalities. For this study, we chose two of those patterns in social human-human communication as reference points: getting acquainted and sharing of emotions (Skowron et al. 2013). When two previously unacquainted people meet for the first time, whether online or face-to-face, they try to reduce uncertainty by “getting to know” one another. They exchange personal information, making themselves known to one another, a process called acquaintance (Altman and Taylor 1973). Another important process at play when two people meet and converse is called social sharing of emotion (Rimé 2009). It consists in exchanging information about what happened to one another, and more specifically about the events which elicited emotional responses: e.g., anger over a recent public transport strike or happiness about one’s latest personal success. Both processes—acquaintance and social sharing of emotion—are core components of everyday conversations, fostering relationship development and maintenance. They have been studied for several decades and can be elicited experimentally.

14.5 Experimental Results

In the following, we give a short overview of the main experimental results for each of the three studies introduced above.

14.5.1 Study 1

Interactions in all four experimental settings (2 × WOZ, 2 × System) were completed by 35 participants (13 female, 22 male), aged between 20 and 50, resulting in 140 interaction logs. After each of the experimental interactions, lasting 5 min, participants were asked the following questions for assessing the conversational system (VH = Virtual Human, the label used for the graphical bartender avatar):

1. Did you find the dialog with the VH to be realistic?
2. How did you enjoy chatting with the VH?
3. Did you find a kind of emotional connection between you and the VH?

The participants provided their ratings on a six-point scale, i.e., from 1 = not at all to 6 = very much. Figure 14.1 presents the aggregated results obtained for the experimental settings with the Affect Bartender and for those with a Wizard-of-Oz.² In all three tasks, the results achieved by the conversational system match those

²In all figures, data are normalized by the number of utterances emitted by a user in a given interaction. Asterisks indicate significant differences at p < 0.05. Error bars indicate ±1 standard error.
obtained for the WOZ. In particular, the correlation coefficients for the aggregated AB and WOZ ratings varied between 0.95 (chatting enjoyment), 0.96 (emotional connection) and 0.97 (dialog realism). All these correlations differ from 0 at a significance level of 0.001. A repeated measures analysis of variance showed no main effect of the setting (AB vs. WOZ) on the three dependent measures (all Fs(1, 34) < 0.50, ps > 0.49). Pairwise comparisons with Bonferroni correction (Holm 1979) confirm the absence of significant differences between the two settings on the perception of dialog realism, chatting enjoyment, and subjective feeling of emotional connection with the system.

**14.5.1.1 Comparisons Between System and WOZ Data**

Comparisons between system and WOZ utterances demonstrate success in conveying interest in feelings and concerns of users, potentially connecting with the user. Specifically, examined with a lexicon-based sentiment classifier (Paltoglou et al. 2010), system utterances were both more positive, and more negative compared to WOZ utterances (repeated measures analysis of variance, Fs(1, 34) > 36.61, ps < 0.001). Simply put, system utterances were more emotionally loaded than WOZ utterances, both in the positive and in the negative orientation. Additionally, when examining utterances with LIWC personal concerns categories, we find that the system generated significantly more words related to work, home, money, religion, and death, compared to the WOZ (Fs(1, 34) > 11.59, ps < 0.01). In short, the system was able to talk more about emotions as well as potential user concerns, in an attempt to relate to users’ feelings and interests.
14.5.2 Study 2

Interactions were completed by 91 participants (33 female, 58 male), aged between 18 and 52, in all three experimental settings resulting in 273 interaction logs.

14.5.2.1 Effects of Affective Profile on System’s Evaluation and Emotional Changes

The affective profile had a series of significant effects on the evaluation of the system and on users’ emotional changes. Detailed results of the analyses performed are presented in Skowron et al. (2011c,a). Concerning evaluation, the affective profile had significant effects on all dependent measures (see Fig. 14.2). The largest effect sizes were found on statements 5 and 6 (positive and negative emotional change, respectively). As expected, affective profiles successfully induced corresponding emotional changes in users, affecting perception of dialog realism and coherence only to a smaller extent.

![Fig. 14.2](image)

Fig. 14.2 Participant’s mean ratings on all dependent variables, of their interactions with the dialog system (DS) with three different affective profiles (positive, neutral, negative) (Skowron et al. 2011c)
14.5.2.2 Effects of Affective Profile on Users’ Interaction Style

Users were equally fast in replying to different affective profiles. Specifically, when analyzing the whole interactions, there were no differences in the participants’ average response time to a number of letters generated. They also used an equal amount of words and utterances in their conversations for all profiles. There were, however, significant differences in word categories used and other linguistic aspects of the text input. Among others, compared with the positive profile, the negative profile elicited, as expected, less assent (e.g., ok, yes, agree) from users, fewer positive emotion words, more anger-related words, and utterances assessed as significantly less positive by the sentiment and ANEW classifiers (see Fig. 14.3). The positive profile, on the other hand, elicited accordingly more positive emoticons, more positive emotion words (e.g., love, nice, sweet), more user statements, and

![Bar chart A](https://example.com/bar_chart_a.png)

![Bar chart B](https://example.com/bar_chart_b.png)

**Fig. 14.3** Valence, arousal, and dominance ratings in user exchanges with the dialog system (DS). Panel (a) shows the mean positive and negative Sentiment Classifier score per condition. Panel (b) shows the mean valence, arousal, and dominance scores based on the ANEW lexicon (Skowron et al. 2011a)
less closed questions to the system. The two latter findings might indicate more information disclosure and less questioning from users towards the positive profile, compared to the negative profile.

14.5.3 Study 3

Interactions were completed by 75 participants (38 female, 37 male), aged between 18 and 32, in all three experimental settings resulting in 225 interaction logs.

14.5.3.1 Effect of Communication Scenario on System’s Evaluation

A multivariate repeated measures analysis of variance showed the expected absence of effect of communication scenario on users’ evaluations of the system (Wilk’s $\lambda = 0.90$, $F(14, 276) = 1.07$, $p = 0.39$). In other terms, participants judged the three communication scenarios (neutral, getting acquainted, and social sharing of emotion) as equally enjoyable, coherent, realistic, emotionally connecting, etc. (univariate tests: all $F$s(2, 144) $< 2.47$, $ps > 0.09$).

14.5.3.2 Effects of Communication Scenario on Users’ Interaction Style

Words and Timing Participants conversed significantly more in the “social sharing of emotion” scenario, despite the time limitation. There was a main multivariate effect of communication scenario on indicators of conversation length (Wilk’s $\lambda = 0.84$, $F(6, 284) = 4.48$, $p < 0.001$). Univariate analysis showed that each length indicator, i.e., character, word, and utterance count, was affected by the communication scenario ($F$s(2, 144) $> 3.03$, $ps \leq 0.05$). Post-hoc comparisons, shown in Table 1 (Skowron et al. 2013), revealed that participants wrote significantly more when sharing an emotional episode.

Confirming these results, we found a main effect of communication scenario on response time ($F(2, 144) = 8.29$, $p < 0.001$). Specifically, participants wrote significantly faster in the social sharing of emotion scenario, compared to the two other conditions.

Dialog Act Classes Participants wrote significantly less statements overall in the social sharing scenario, and significantly more statements about ordering food and drinks in the neutral scenario ($F$s(2, 144) $> 5.63$, $p < 0.01$).
14.5.3.3 Effect of Communication Scenario on User’s Expression of Affective States

**Sentiment Classifier, and ANEW Lexicon** There was a main effect of communication scenario on both positive sentiment score ($F(2, 144) = 5.59, p < 0.001$) and ANEW ratings (Wilk’s $\lambda = 0.60$, $F(6, 284) = 13.77, p < 0.001$). As depicted in Fig. 14.4, participants wrote significantly less positive, less arousing, and less dominant utterances in the social sharing of emotion scenario; note the negative context of the sharing introduced by the system’s “personal experience”. The largest effect size is found for ANEW valence ratings ($\eta^2_p = 0.22$) first indication of the

![Fig. 14.4 Study 3: Effect on expressions of affective states and sentiment. Communication scenarios: BASE–neutral communication scenario, ACQ–getting acquainted scenario, SHARE–social sharing of emotion scenario. Panel A shows the mean positive and negative sentiment classifier score; panel B shows the mean valence, arousal, and dominance scores based on the ANEW lexicon (Skowron et al. 2013)](image-url)
successful creation of a social sharing of emotion situation, where participants shared a negative experience, which did not happen in the two other conditions.

14.5.4 Summary of Experimental Results

In the presented experiments, a variant of IAS, the Affective Dialog System, was applied to study communication processes and the impact of affect and social processes on users in online interactions. The experiment specifically focused on challenges of structuring emotional interactions in open-domain dialogs limited to the textual modality, and on investigating the impact of such interactions on users. The first study, conducted in a WOZ setting, validated the system’s ability, on par with a human operator regarding realistic and enjoyable dialog as well as for establishing an emotional connection with users in a short interaction. The second study confirmed a significant influence of the system’s affective profile on users’ perception of a conversational partner, on the changes of emotional states reported by the users, and on the textual expressions of affective states. The third study proved a successful application of the dialog system for eliciting social sharing of emotion and for realizing a communication scenario of ‘getting acquainted’. The experimental results show the impact of the realized communication scenario on the communication style of users, their expressions of affective states and self-reported emotional changes experienced when interacting with the system.

The reported experiments applied realizations of IASs in dyadic and triadic settings. The next step in our research, the application of IASs to group interactions creates a new set of challenges related with, e.g., simultaneous communication with multiple users (beyond triadic settings), capacities to interact in a way which intentionally follows or violates the typical communication patterns of members of a particular online community.

14.6 Scaling up to Multiple Users Environments

To date, different realizations of IASs have been applied in experimental setups where the system interacted in dyadic and triadic settings. The purpose of these experiments was to evaluate the system’s ability to participate in a coherent dialog, and to establish and maintain an emotional connection with users. These experiments shed light on how affective system profiles and fine-grained communication scenarios impact the self-reported emotional changes of users, their communication style, and their textual expressions of affect. The next step, the application of IASs to group interactions created a new set of challenges related with, e.g., simultaneous communication with multiple users (beyond triadic setting), capacities to interact in a way which intentionally follows or violates the typical communication patterns of members of a particular online community. This included factors such as the affective dimension of interactions, and the ability to observe such a behaviour in
other participants. These functionalities impact the system’s capacity to generate consistent or intentionally inconsistent interactive behaviour, the required affective coherence and the event-dependent adaptation of its interaction patterns to other members in a group. Simultaneously, the system needs to perceive, represent and model discussions and emotional exchanges, at the individual and group levels. This prerequisites abilities to: (i) predict the possible outcomes of the observed group dynamics, (ii) simulate the effects of system’s interactions with individuals or a group, and (iii) to assess the real effect of its interventions and to correspondingly update the used models.

To address the requirements related with transferring IASs from dyadic and triadic to multi-user interaction settings, the proposed approach integrated experience gathered in conducted experiments with insights from a wide range of studies on the role of emotions in online communication: psychological studies and experiments on perception and generation of emotionally charged online content (Kappas et al. 2010; Küster et al. 2011; Küster and Kappas 2014), event-based network discourse analysis (Hillmann and Trier 2012), agent based models of emotions (Schweitzer and Garcia 2010; Garas et al. 2012), valence trends (Chmiel et al. 2011a,b) and agent based model on bipartite networks (Mitrović et al. 2011; Gligorijević et al. 2013), presented in the earlier chapters. For more extensive presentation of the models, their transfer and integration in IASs refer to Skowron and Rank (2012) and Rank et al. (2013).

14.6.1 Role of Simulations in IAS

The role of agent-based modelling and simulation in IASs is twofold: (i) to provide cues on information provided to participants, and (ii) to serve as decision support. For example, based on a request from a particular e-community the tools can support it with analysis on affective dimension of their interactions and suggest ways for counter-acting negative tendencies observed in a group, e.g., growing hostility between members, or a decrease of cooperation. Simulation results can also indicate targets for interventions. This entails several requirements that relate to the results of simulation runs, runtime characteristics and the adaptability of the simulation based on data collected during previous interactions. Here, several questions relevant as a potential input for the systems’ decision making mechanisms were identified (Rank 2010):

- Which individual in a group will be most likely to provide an accurate response to probing about the group’s emotional state, and which one will be most reliable?
- What influence can individuals have on the evolution of the collective emotions in an e-community, and which of the specific participants is likely to have the biggest influence?
- Can potential escalations, both in the negative and in the positive direction, be detected early on?
- What influence will a specific intervention of the system have at the current moment, and which style of intervention is most effective?
Running a simulation on demand to query about the above questions adds the requirement of timely, or possibly anytime, responses but also the need to parameterise the simulation to promptly respond to the current state of an online community, ideally using the recorded history as input. An important part of the decision-making structures of IASs is the modelling of conversation participants. This component of the agent control structure is analogous to adaptive user modelling in standard HCI: the system initially has a default model of the interaction partner, adapts it over time, and complements missing information based on the knowledge derived from interaction events. In the case of multi-user environments, this includes modelling several participants, simplifying the employed models and abstracting from specific individuals. The modelling eventually serves the purpose of deciding on utterance selection, utterance modification, timing of utterances, and the selection of conversational partners in multi-user environments. As such, the main questions that modelling efforts helps to answer for the purposes of affective interactive systems are, from general to specific:

- What potential influence will certain interventions have on the collective state of an online community?
- What is the influence of particular interventions on the future development of a specific group?
- What type of intervention (affective charge, topic, timing) will have which effect?
- What relation does a particular individual have to the state of a specific group?
- Which intervention is most appropriate when addressing a particular individual of a specific group?

### 14.6.2 Input from Theoretical Modelling and Analysis

The questions presented above relate the decisions of individual agents to the collective emotions of the group, and require a dedicated approach that can deal with the relation between individual and collective levels. We extended our simulation of emotions in IAS with a model for the real-time evolution of emotional states, based on the modeling framework presented in Chap. 10 (Schweitzer and Garcia 2010). Within this framework, we can define the dynamics of individual emotions based on the concept of Brownian agents (Schweitzer 2003), integrating empirical data with analytical results that show the emergence of collective emotions in multi-user environments.

Applying this framework, we can integrate previous findings from very different online communities, including product reviews communities (Garcia and Schweitzer 2011), social networking sites (Šuvakov et al. 2012), and blogs (Mitrović and Tadić 2012). A particularly relevant application for IASs is the model for emotional persistence in online chatting communities presented in Chap. 10 (Garas et al. 2012), which provides a description of individual emotion dynamics for the users of a real-time group discussion.
14.6.2.1 Activity Patterns in Time

The individual dynamics of emotions in our model are based on the empirical analysis of user interaction in real general discussions in Internet Relay Chat (IRC).\(^3\) We analyzed the time component of the interaction in the chatroom through the distribution of time intervals between messages of the same user, shown in Fig. 14.5.

Similarly to the case of short message service communication (Wu et al. 2010), this distribution has two modes, one corresponding to the time intervals between bursts of interaction (light gray), and another one containing the time intervals within an interaction burst (dark gray). The latter is characterized by a power-law distribution of the form \( P(\Delta t) \sim \Delta t^{-\alpha} \) for \( \alpha = 1.54 \), where the former is better explained by a lognormal distribution. To provide the most realistic behavior possible, the agent actions in our model are chosen to emulate this empirical distribution. For the case of IASs, every reaction has an additional delay \( \Delta t \), sampled from the power-law regime of the distribution, i.e. the inter-activity times associated with group discussions.

The addition of this stochastic delay changes the properties of the IAS in a substantial manner that is commonly ignored in this kind of system. A usual design decision driven by intuition is assuming that the time between messages of an agent can be sampled from a distribution with a given mean (e.g. normal). On the other hand, the empirical distribution of time intervals between messages in a group IRC discussion is very different, following a power-law of exponent 1.54, which does not have a finite first moment. Therefore, empirical evidence contradicts the assumption mentioned before, and consequently the model used in IASs has a time behavior closer to the real patterns in online interaction.

\(^3\)The analysed data-set included 2.5 million posts acquired from EFNET IRC chats: http://www.efnet.org, covering a range of topics including music, casual chats, business, sports, politics, computers, operating systems and specific computer programs.
14.6.2.2 Emotional Persistence

Sentiment analysis techniques like the ones described in Chap. 6 can extract the emotional content of short chat-room messages, giving a value of polarity per message. Using this method a sequence of messages would be represented as an ordered set of ones for positive messages, zeroes for neutral messages, and minus ones for negative messages. Assuming a sufficiently large data set, these kinds of sequences can be processed with tools from statistical physics, revealing properties of the emotions expressed at individual and group levels. One of these techniques is called Detrended Fluctuation Analysis (Peng et al. 1994), which can be used to calculate the Hurst exponent (Hurst 1951; Rybski et al. 2009), a measure of persistence in a time series. This persistence is reflected in the time series as consistent fluctuations around its mean, revealing states of consistent biases towards positive or negative emotions.

Persistence can be calculated for an individual user, when processing the sequence of emotional expressions of that user, or at the discussion level, when computing it over all the messages in a group discussion. If a user behaves in a fully random way, with fixed probabilities of each sentiment polarity but no memory, it would have a persistence of 0.5 and a mean emotional expression according to the given probabilities. A user with some emotional momentum, who tends to express emotional states similar to previous ones, would have a persistence above 0.5. This case of persistent users is the most common one, which can be appreciated in the kernel density plot of Fig. 14.6 in combination with the distribution of mean emotional expression. Most of the users show a bias towards positive emotional expression, with persistence in the way these emotions are expressed though the discussion.

Fig. 14.6 Kernel density plot of the distribution of individual persistences and mean emotional expression, for all users with more than 100 messages in the observed period (Garas et al. 2012)
We measure the collective emotions in the whole discussion by computing the persistence over all the messages in the group chat. The empirical analysis of 20 IRC channels (Garas et al. 2012) showed persistence values above 0.5 for all channels, revealing the existence of collective emotional states at the aggregated level. This appears in the discussion as fluctuations around the mean emotional content, which is also in general biased towards positive emotions. This collective positivity and persistence shows that users influence each other through their expression of emotions, and that certain rules of emotional expression are present even in the anonymous and ephemeral discussions of IRC chat-rooms.

14.6.2.3 Brownian Agents Model

Although a persistence pattern for an individual user can be easily modeled with a biased random walk, the emergence of such collective persistence does not trivially follow from individual emotional expression. If we design a model which only ensures persistence in the behavior of an individual, the combination of expressions of many agents would show no persistence at the group level. Our aim is to provide realistic emotion dynamics, useful for IASs, in which collective persistence appears in a discussion with many agents, which we reproduce with the Brownian agents model.

The agent dynamics of this model are outlined in Chap. 10, where the arousal dynamics is given by the time interval distribution shown in Fig. 14.5. Agents in this model have an externally observable variable $s$, which corresponds to the messages posted by the IAS in the collective discussion. This variable of expression is activated according to the arousal, and its value is given by the internal valence of the agent. The communication in the group discussion is modeled through an information field, which aggregates the emotional expression of all the agents. This field influences the dynamics of individual valences, creating a herding effect that reproduces collective persistence from the interaction of many individuals. Simulations of this model show how mean emotional expression and persistence depend on the parameter values (Garas et al. 2012), allowing us to simulate discussions that reproduce the observed group behavior. In terms of individual dynamics, the emotional profile of an agent can be sampled from the distribution shown in Fig. 14.6, choosing values of emotional tendency and persistence for the agent. Such Data-driven simulations with the corresponding parameter values lead to the desired individual and collective persistence effects.

14.6.2.4 Model Extensions

The Brownian agents model outlined here is a first approximation to the dynamics of emotions in e-communities that can be improved and extended with additional features within the same modelling framework. The assumptions followed by models like the one explained above are currently being tested in experimental
setups like the ones introduced in Chap. 5, which already revealed the patterns of influence of online interaction in internal emotions. In this model, interaction takes place in a publicly accessible discussion in which messages are not directed to individual users. Some of the questions listed in Sect. 14.6.1 aim at driving IASs in a way such that they can interact with individual conversation participants, by previous analysis of their position and influence in the overall discussion. Some models within this framework include a network component that precisely deals with this kind of user heterogeneity, as explained in Chap. 11 (Mitrović and Tadić 2012; Šuvakov et al. 2012).

In real-time interaction, the network mapping approach (Gligorijević et al. 2012a) analysis accounts for the properties of activity patterns and underlying network topologies characteristic for various types of users, including those identified as important/influential in a given online interaction environment. In IASs, the spanning trees analysis can be specifically applied to analyze: (i) the user’s activity, i.e.: by the creation and analysis of evolution of the network links, e.g., positive, negative, and (ii) users collective behaviour patterns. As demonstrated in high-resolution analyses of user-to-user communication in IRC channels, only certain links survive over 1 day period and support a particular type of network structure. Based on this observation, the presented system realizations and application scenarios presented in Sect. 14.2.1 are primarily targeted at serving on-demand information requests of e-communities or individuals, i.e., establishing a relatively short-time direct communication links. As experimental evidence demonstrates, such direct, limited in time interactions with users, also contribute to an overall higher level of the perceived realism, dialog coherence, the feeling of an emotional connection and chatting enjoyment (Skowron and Rank 2014). In practical application and deployment scenarios, this often translates to a higher acceptance rate of interactive systems in online communities.

### 14.7 Ethical Considerations

Endowing interactive systems with the capability to model complex affective processes and to account for affective dynamics in networks opens up a range of applications for assisting e-communities; however, it also allows for malicious misuse. “Short-sighted” commercial, political, or ideological causes of individuals or groups could lead to the use of the technology for generating, amplifying, or reducing naturally occurring collective emotions. Such misuse could result in an “arms race” between individuals or groups with competing interests resulting in a decrease of trustworthiness of online communication channels.

This leads to a question on how these technologies should be developed, tested, and applied in real-world e-communities. The aim of such a debate is early

---

anticipation of risk and identification of likely consequences, both beneficial and harmful related to field of research and its real-world application in the line of Novotny et al. (2003) who postulate: “against a background of inherent uncertainty about the future state of knowledge (and almost everything else) from which, of course, scientific potential is derived, it is necessary to reach beyond the knowable context of application to the unknowable context of implication. Here knowledge producers have to reach out and anticipate reflexively the implications of research processes”. The discussion further requires public dialog, broader than user involvement, while the role of researchers in this aspects shall relate to analysis and explanation of dystopian and utopian scenarios, identification and development of strategies to avoid likely harmful scenarios, and promote likely beneficial ones (Prescott 2013).

In a scientific setting, the acquisition of data and the engagement of experimental subjects should always be accompanied with ethical considerations and the adherence to a code-of-conduct that is intended to protect the rights of the individuals involved. In our opinion, the usage of such technologies in non-laboratory settings requires the consideration of rules of engagement similar to a code-of-conduct that specifies the scope of interventions and ascertains an equivalent of informed consent both from individuals and from e-communities as a whole as far as possible. The new opportunities offered need to be balanced with the potential threats of mis- and abuse. As a minimal rule of engagement, we assume open identification of such systems when interacting with users so that they are informed about the presence of this influence.

In a range of application scenarios, a prerequisite for a successful application of such interactive systems could be also the ability to adjust (or at least foresee the outcomes of an intentional violation of) one’s communication behavior or affective stance according to: the overall mood detected in a group; individuals’ preferences to various entities or fellow participants; the established or evolving “social norms”; or dynamic changes in a hierarchy of interaction patterns of users. The ability to follow, i.e., perceive, model and adhere to social norms used in a given e-community could be treated as an additional prerequisite for the deployment of IASs to the real-world online groups.

Such a convention can be advantageous for application development too, since the acceptance of artificial systems in online communities depends on the perceived benefits and costs associated with the systems’ activities. For the purpose of IASs, this dependency is translated in the proposed systems directly into their communications policy: the system communicates directly with users in situations where the relevance of a potential contribution, i.e. added informational or affective value—contribution value, is estimated to be sufficient with high confidence. These estimates are based on user and group models for the impact of specific content. Additional action costs are associated with interaction scenarios involving larger numbers of users or, progressively more costly actions, which include the

\footnote{In an absence of other, higher level design principles.}
community as a whole, e.g. where posts need to be emitted to all users at the same time. This is also reflected in an IAS’s ability to perceive and model the affective stance of individuals and groups towards itself and to adapt in order to improve its perceived usefulness.

Further specific areas for discussion also include question on how e-communities could control the extent to which artificial systems influence their communication patterns including both the exchange of information, an area that has been studied for quite some time, as well as the affective dimension, a relatively new area with a potentially stronger impact.

A positive example for an emerging application scenario is the application of IASs for analysing and acquiring information on online dynamics in interaction environments where large number of young user are involved, e.g., detecting the formations of affective dynamics specific for cyber-bullying or monitoring for other anomalies which can be observed by analysing and modelling of affective communication patterns, or changes in the underlying network and communication structures.

### 14.8 Conclusions and Outlook

The rapid expansion and evolution of ICT-mediated communication channels requires methodologies tailored for studying collective emotions in e-communities and the development of tools that can support the evolving needs of these communities. Agents’ online interactions, mediated over a range of communication channels, have the potential to influence the processes of formation and evolution of collective emotions. Emotional effects can rapidly spread across the virtual realm regardless of the physical location of peers or their offline-world group identification. Consequently, even a single agent’s actions and interactions with other members of e-communities can have a strong impact on collective emotions there. This applies both for the human and artificial agents and the presented experiments demonstrate that artificial systems can be useful for studying this impact experimentally and in an interactive way that complements other forms of study.

The system applications presented in this chapter focus on text-based, real-time system-user communication aimed at the detection, acquisition and modeling of users’ affective states. The scope of communication with the system was not limited to a specific domain or one particular e-community. Clearly, interactive systems like these fall short in their conversational abilities in comparison with humans, in particular in longer lasting interaction scenarios, potentially integrating open- and closed-domain dialog and discourse processing. However, as the experimental evidence demonstrates for relatively short communication scenarios, the system could match the WOZ results in terms of dialog realism, chatting enjoyment and the ability to establish an emotional connection with subjects. The analysis of activity patterns and affective dimensions of users’ communication in multi-user
environments revealed that the majority of links are established only temporarily and are used primarily to exchange relevant information, to share or respond to a sentiment expressed. These findings support the application scenarios where an IAS establishes communication links similarly to human counterparts: briefly and on demand. Further, the system should be able to predict likely outcomes of its own behaviour, either following or intentionally violating established or evolving social norms, communication conventions, or the communicative style in a group, in order to adjust its behavior accordingly. The modelling of self, including the affective reception of content provided, both in 1-on-1 and multi-users interactions, allows for adjustments of thresholds used for calculating action benefits and costs. Correspondingly, as a starting point, we proposed application scenarios where systems communicate directly with users in situations where high confidence scores for a potential contribution’s relevance, i.e. added informational or affective value—contribution value, can be foreseen. Estimates are based on outcomes of simulating the reception of a specific content by a particular individual or a group. Further, additional action costs should be accounted for in interaction scenarios where posts need to be emitted to a large number of participants or to the whole e-community. In this chapter, we also presented design choices and specific models for applying agent-based simulation as part of the decision mechanism in the presented systems.

The development and application of Interactive Affective Systems relates to a range of scientific challenges from affect detection, modelling, and generation, to the modelling of interactive, affective and social dimensions in individuals and groups over time. In parallel, models need to account for the dynamics and complexity of underlying social networks as well as of aspects of human-computer interaction. The type of system introduced here can act, e.g., as e-community advisers aiming at increased cooperation in a group, counteracting growing hostilities, and yielding measurable social benefits. However, there are also application scenarios where such systems can potentially be abused for achieving adverse effects. As the study and application of artificial affective systems is expanding, there is also an urgent need to define suitable ethical rule-sets, i.e. code of conduct for real-world use in online communities.

References


