Let's Talk! Socially Intelligent Agents for Language Conversation Training

Helmut Prendinger and Mitsuru Ishizuka
Department of Information and Communication Engineering
School of Engineering, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
{helmut,ishizuka}@miv.t.u-tokyo.ac.jp

Abstract

This paper promotes socially intelligent animated agents for the pedagogical task of English conversation training for native speakers of Japanese. Since student-agent conversations are realized as role-playing interactions, strong requirements are imposed on the agents' affective and social abilities. As a novel feature, *social role awareness* is introduced to animated conversational agents, that are by now strong affective reasoners, but otherwise often lack the social competence observed with humans. In particular, humans may easily adjust their behavior depending on their respective role in a social setting, whereas their synthetic pendants tend to be driven mostly by attitudes, emotions, and personality. Our main contribution is the incorporation of 'social filter programs' to mental models of animated agents. Those programs may qualify an agent's expression of its emotional state by the social context, thereby enhancing the agent's believability as a conversational partner. Our implemented system is entirely web-based and demonstrates socially aware animated agents in a virtual coffee shop environment. Preliminary experiments with our conversation system indicate that animated agents with social competence and affective behavior are considered as interesting training companions for language students.

Keywords

Social role awareness, social dimension in communication, believability, affective reasoning, emotion expression

I. Introduction and Motivation

We can recently notice a shift of interest from believable, life-like agents to a new generation of agents that are characterized as 'socially intelligent' (Dautenhahn [14]). Rather than being completely different from believable agents, the socially intelligent successors are intended to perform more natural and robust in interactions that are explicitly embedded in a social context with associated social goals, such as establishing and maintaining social relationships (Dautenhahn [11], Bickmore and Cassell [4]). However, the added value is crucial since, as demonstrated experimentally by Reeves and Nass [37], people treat computers as if they were veritable social actors.

Ever since Bates and Reilly promoted *believable* agents in their 'Oz project' [3], [25], there have been continued efforts to give animated agents the illusion of life. By now, it is widely accepted that emotion expression and personality are key components of believable agents (e.g., Rousseau and Hayes-Roth [39]). Moreover, Cassell and her co-workers [8] provided convincing evidence of the importance of non-verbal 'embodied' conversational behavior for believable agents. On the other hand, *life-likeness* is often more concerned with the realism of animated (or synthetic) agents (Cassell [7]).

In the case of *socially intelligent* agents, opinions regarding the core issues are more diverse, possibly because the field is highly interdisciplinary [13]. As a guideline, we report on Dautenhahn's [12] four assumptions ('working hypotheses') underlying her work on social agents: (i) intelligence is linked to a body, (ii) the body is adapted to its embedding environment, (iii) intelligence can only be studied when considering the interaction between body and environment, and (iv) intelligent agents are social entities. She defines *social intelligence* generally as

> [...] the individual's capability to develop and manage relationships between individualized, autobiographic agents which, by means of communication, build up shared social interaction structures which help to integrate and manage the individual's basic ('selfish') interests in relationship to the interests of the social system at the next higher level.

In particular, Dautenhahn [12] argues that social intelligence in *communication* is based on synchronization processes that involve both external coordination of behaviors (e.g., speech acts) and internal phenomena such as empathic understanding which might lead to social understanding and social learning.

In this paper, we will take a particular stance on socially intelligent agents. While our primary interest is agents' believability as conversational partners, we will argue that purely emotion and personality based agent behavior falls short of 'social robustness', as it does not account for the social context in which a conversation is embedded. E.g., a sales agent might be convincing (believable) in the role of a salesperson, but not in the role of a colleague or employee. Humans, on the other hand, are always aware of the roles they play in a certain social setting and typically behave accordingly. This important feature of human-human communication will be called *social role awareness* and constitutes our contribution to the social intelligence of agents. As an example of social role awareness, consider an aggressive sales agent interacting with a customer, which is angry at the customer because he or she is asking detailed questions about the product. The agent will presumably use polite language and hide its anger, as it is aware of the role-specific behavioral restrictions applicable to a sales conversation. On the other hand, imagine the sales agent interacts with a colleague (agent). In this social setting, we would expect that the sales agent behaves according to its personality traits (in our case, aggressive). A basic assumption of our work is that agents' behavior cannot be generated (or understood) by considering personality, attitudes, and emotions alone, but has to integrate social role awareness as an essential component.

We recently started a project with the aim to employ animated agents for the pedagogical task of *language conversation training* (Prendinger and Ishizuka [36]). Specifically, the animated agent approach is used to improve English conversation skills of native speakers of Japanese. The setup of the project is guided by two methodological assumptions. First, we rely on empirical results on animated pedagogical agents that demonstrate their motivational role in the learning context, in addition to the possibility of increased learning effectiveness. Lester *et al.* [23] show the affective impact of animated agents on student's learning experiences, that revealed the so-called 'persona effect', i.e., the mere presence of an animated character may increase the user's arousal and motivation to perform a task well. Second, as interactions between language students and agents are implemented as role-playing dramas and games, strong requirements are imposed on the agents' affective and social abilities. In order to come across as believable and interesting conversational partners, the agents have to show coherent emotional responses and social competence.

Social reasoning, here understood as an agent's reasoning about the social context, will be blended with a rather standard theory of reasoning about emotions (Ortony *et al.* [34], [33]) and a simple model of personality (Moffat [28]). We employ Moulin and Rousseau's [30] approach to modeling and simulating conversations, which provides a rich framework for many aspects of inter-agent communication. Currently, the programmable interface of the Microsoft Agent package [27] is used to run our role-playing scenarios. Although the off-the-shelf agent characters available for this package are quite restricted in the number of behaviors, the package comes ready with a speech recognizer and text-to-speech engine that allow client-side execution in a web browser. It should be noticed, however, that our technology for socially intelligent agents is independent of the agent system, i.e., it could also be integrated to the significantly more realistic (life-like) agents, as the 'embodied conversational agents' discussed by Cassell [7].

The main contribution of our work is twofold. *First*, we develop a principled way to model agents that can be considered as socially intelligent in their conversational behavior. We focus on subtle variations in agents' emotional responses, depending on their affective state, personality, and the social setting. *Second*, we describe a web-based implementation of socially aware agents and give a preliminary evaluation of their effectiveness as conversational partners for role-playing interactions. A major limitation of our work is that the evolution of social relationships as a consequence of social interaction is insufficiently considered. However, future improvements of social dynamics are outlined.

The rest of this paper is organized as follows. The next section discusses related work on believable social agents and on language training. In Section 3, we describe a framework for modeling and simulating conversations among multiple agents. In Section 4, we first introduce an affective reasoner for reasoning about emotions and a rudimentary model of personality. After that, we explain the basic notions underlying social reasoning. Moreover, a simple model for emotion expression is described. Section 5 is the core part of our paper. Here we argue that reasoning about emotions and personality is not sufficient to achieve believable emotion expression, and introduce social filter programs that function as a filter between an agent's affective state and the agent's emotion expression. In Section 6, we first explain the web technology employed to run the animated agents, and then we illustrate our approach by example runs of a Cybercafé-style role-playing drama. We also report about early feedback from students using our conversation training system. Section 7 discusses the paper and suggests further extensions and refinements. Finally, we summarize the paper in Section 8.

II. Related Work

Several research groups have addressed the problem of socially intelligent agents in the framework of multi-agent systems (e.g., Jennings [22], Castelfranchi [9]). On the other hand, there are relatively few researchers who focus on social role related behavior from the perspective of the believability of animated agents. After a brief discussion on related work on intelligent agents, we will mention two commercial software tools that are used to improve English conversation skills.

A. Socially Intelligent Agents and Believability

Hayes-Roth and her co-workers [21] study role-specific behavior in the context of characters (animated agents) that function as actors in a master-servant scenario. Social roles are defined by behaviors that represent a character's status, e.g., high status is effected through a quiet manner and ways of talking that forbid interrupting. A character's believability in a specific role is justified by adopting guidelines from the literature on the art of drama, rather than by exploiting a social-psychological model. Thus, no attention is paid to a character's representation of its role. Rousseau and Hayes-Roth [39], however, propose an elaborate social-psychological model that considers personality, emotions (moods), and attitudes. Rules are defined that allow to select appropriate behavior depending on a character's personality and attitudes, thereby enhancing the character's believability. Interestingly, the social roles discussed in [21] are only of marginal importance in this model. By contrast, we believe that an agent's awareness of its social role is equally important for action selection, and may even overrule the influence of personality.

Walker *et al.* [42] promote linguistic style as a key aspect for believable agents. Linguistic choices are seen as realizations of an agent's personality, and subject to social variables such as social distance between agents or the power one agent has over another. The *Linguistic Style Improvisation* (LSI) framework is based on speech acts theory and a theory of linguistic social interaction. As a theory of social conversation, LSI is clearly superior to our approach, which does not provide a formalization of speech acts and uses a simpler algorithm to decide the linguistic style of utterances. However, we will provide a subtler account of the interaction between an agent's emotional state, personality, social context, and emotion expression.

Gratch [18] introduces 'social control programs' on top of a general purpose planning system. In this system, plan generation and execution are biased by the characteristics of the social context. A so-called 'personality GUI' contains the agent's goals, its social status, etiquette, (in)dependence, and attitudes towards other agents. Besides those static features of the agent's mental state, Gratch [18] introduces dynamic features of the social context such as the communicative state, the plan state, and the agent's emotional state. Social rules encode commonsense rules of social interaction, e.g., help a friend or avoid that some other agent interferes with your plan. For the rather rigid organizational setting in which military commander agents operate, Gratch and Hill [19] introduce social concepts similar to ours. The main difference to our work is that we place social control programs at the interface of the module that reasons about emotion and the module that renders the emotional state to actual behavior. By considering the social context, we aim to achieve believable emotion expression rather than generate socially adjusted plans.

Guye-Vuillème and Thalmann recently started to work on an architecture for believable social agents which is based on four sociological concepts: social norms, values, world view, and social role (see [20] for an abstract).

Finally, the research closely relevant to ours is done by Moulin and collaborators [29], [6], [30], who explicitly integrate the social level to their model of conversation among multiple agents. Their model will be discussed in more detail below.

B. Language Conversation Training

Serving a similar purpose as our conversation training system, *Native World* [32], a commercial software tool, allows users to practice English by interacting with native speakers of English displayed in video clips. After each dialogue move of the user, via speech recognition of a sentence shown on the screen, a new clip is loaded from a CD-ROM. Although recent empirical studies demonstrated the effectiveness of video clip based personas (McBreen *et al.* [26]), this tool assumes that video clips are prepared for all dialogue moves. By contrast, our animated agents based approach gives us far more flexibility in generating a variety of socially situated responses. Moreover, we envision client-side execution in a web browser as the learning environment.

Virtual Personalities [41] introduced a virtual human called *Sylvie* (a so-called 'Verbot'), which is an animated face that can be used for Eliza-style conversations, i.e., a the user can type in an English sentence (typically a question), and Sylvie will output an answer via a text-to-speech engine. Depending on how elaborate Sylvie is scripted, her answers are more or less appropriate (or 'intelligent'). Although the user is not confined to pre-defined dialogue contributions, it is hard to achieve natural conversation with Sylvie since on many issues, she will not be able to answer satisfactorily.

III. Basics of Modeling and Simulating Conversations

A conversation is often seen as an activity where multiple (locutor-)agents participate and communicate through multiple channels, such as verbal utterances, gestures and facial display. Each agent has its own goals and will try to influence other participants' mental states (beliefs, goals, desires, intentions, emotions, etc.). Moulin and Rousseau [30] distinguish three levels of communication:

- At the *communication level* agents perform activities related to communication maintenance and turn-taking.
- At the *conceptual level* agents transfer concepts.
- At the *social level* agents manage and respect the social relationships that hold between agents.

Our system basically integrates the second and third level, and a very restricted form of the first level. The communicative level implements conversational features of human-human communication, as proposed by Cassell and Thórisson [8]. At the conceptual level, information is passed from one agent to other agents as a (simplified) symbolic representation of the utterance, e.g., if an agent orders tea, this is simply represented as *order_tea*. According to an agent's role in the social context, the social level puts behavioral constraints on agents' actions and emotion expression (Moulin [29]). As described in the Introduction, we consider an agent's ability to respect behavioral constraints and communicative conventions as a key feature for its social intelligence. This issue will be discussed in detail below.

As an example, consider an agent called 'Al' playing the role of a customer and an agent called 'James' in the role of a waiter. Al orders tea from James by saying "May I order tea please?". The corresponding *communicative act* might be formalized as

$$\mathit{com\_act(al,james,order\_tea,polite,happiness,s0)}$$

where the argument 'polite' is an qualitative evaluation of the linguistic style (LS) of the utterance, the argument 'happiness' refers to Al's emotion expression, and $s0$ denotes the situation in which the utterance takes place.

As in [30], we assume that a conversation is governed by

- a *conversational manager* that maintains a model of the conversation, and
- an *environmental manager* that simulates the environment in which the agents are embedded.

For simplicity, we assume that the conversational manager operates on a shared knowledge base that is visible to all agents participating in the conversation (except for the user). It stores all concepts transferred during the conversation by updating the knowledge base with

$$com\_act(S,H,C,LS,E,Sit)$$

facts. The resulting 'model' of the conversation will eventually be substituted by a less simple-minded conversation model incorporating a formalization of speech acts (as, e.g., in Moulin and Rousseau [30]). Moreover, the conversational manager maintains a simple form of turn-taking management, by assigning agents to take turns based on their personality traits. E.g., if James is an extrovert waiter, he would tend to start a conversation with a customer, which is formalized as "*initiative(james,extrovert,take)*".

The environmental manager simulates the world that agents inhabit and updates its (shared) knowledge base with consequences of their actions. E.g., if the agent character Al got his tea in situation $s5$, this will be stored as "*holds(al,has_tea,s5)*". The characteristics of the environment are encoded by a set of facts and rules. Situation calculus is used to describe and reason about change in the environment (Reiter [38]).

IV. Mental Models of Socially Intelligent Agents

Each agent involved in a conversation is assumed to have its own mental model. A mental model may contain different kinds of entities, including world knowledge (beliefs), affective states (emotions, personality traits, attitudes), goals and plans. In this section, we will focus on the core components underlying socially intelligent reasoning: affective reasoning, i.e., reasoning about emotions, personality, and the social context. Furthermore, we will discuss the output of social reasoning—the agent's emotion expression—which is the expression of its affective state qualified by its personality and the social setting in which the communication takes place.

Our assumptions about the agent architecture are fairly simple. It is generally accepted that an affective agent architecture should at least distinguish between 'primary' emotions or (emotional) *reflexes*, and 'seondary' emotions that presume the agent's *deliberation* about its beliefs, goals, and attitudes (Picard [35]). We will propose an emotion model [34] that certainly encodes the deliberative layer. However, the reactive layer might be easily implemented by a set of condition-action rules [40] (at least conceptually). Those rules will fire immediately when the respective conditions are met and output an emotional reaction. In addition, other researchers (e.g., Allen [1]) propose a *meta-management* layer that is responsible for an agent's adaptability by monitoring and controlling reactive and deliberative mechanisms. Currently, our model does not provide means for meta-management.

An interesting aspect of Allen's work [1] is that he considers all 'higher-level' mental concepts such as beliefs, intentions, emotions, and personality as (motivational) *control states* that he defines as "information-bearing representations of an information processing control system". In other words, if a mental concept might function as a predictor of behavior, it can be considered as a control state. E.g., if we can say "she does this because she has a friendly personality", without referring to other of the observed agent's mental concepts (e.g., intentions), it is a good indicator that 'personality' is a control state. A mental concept *not* considered by Allen [1] is social role awareness. We will argue that the social role played by an agent in a certain (social) context may serve as a reliable predictor, and hence should be considered as a control state as well.

A. Affective Reasoning

Many systems that reason about emotion, so-called *affective reasoners*, derive from the influential 'cognitive appraisal for emotions' model of Ortony, Clore, and Collins [34], also known as the OCC model (e.g., Elliott [16], O'Rorke and Ortony [33]). Here, emotions (more precisely, emotion *types*) are seen as an agent's valenced reactions to events, other agents' actions, and objects, qualified by the agent's goals, standards, and preferences. The OCC model groups emotion types according to cognitive eliciting conditions, and labels them by a word or phrase, such as *joy*, *happy-for*, or *angry-at*. In total, twenty-two classes of eliciting conditions are identified. The OCC model can be represented by a set of rules and might thus be seen as a model of (emotion) *causation* that allows to reason about emotion [33]. By way of example, we give the definitions of the emotion types *happy-for* and *angry-at*:

- **Emotion type *happy-for*:** an agent $L1$ is in the emotional state *happy-for* an agent $L2$ in situation S **if**
  - $L1$ experiences joy over a state-of-affairs F presumed to be desirable for $L2$ in S.
- **Emotion type *angry-at*:** an agent $L1$ is in the emotional state *angry-at* another agent $L2$ about action A in situation S **if**
  - agent $L2$ performed action A prior to S **and**
  - action A causes a state-of-affairs F to hold in S **and**
  - agent $L1$ wants the opposite of F to hold in S **and**
  - $L1$ considers action A as blameworthy.

Affective reasoners based on the OCC model can be characterized as theories of 'cold' emotions, where agents simply derive an emotional state based on a cognitive reasoning process and in no way 'have' emotions.

B. Personality

Personality traits are typically characterized by patterns of thought, attitude, and behavior that are permanent or at least change very slowly (e.g., Moffat [28]). Consequently, believable agents should be *consistent* in their behavior (Rousseau and Hayes-Roth [39]). For simplicity, we consider only two dimensions of personality, which seem crucial for social interaction (e.g., André *et al.* [2]).

- *Extraversion* refers to an agent's tendency to take action: sociable, active, talkative, optimistic. We consider the values: 'outgoing', 'neutral', and 'introverted'.
- *Agreeableness* refers to an agent's disposition to be sympathetic: good-natured, helpful, forgiving. We consider the values: 'friendly', 'indifferent', and 'unfriendly'.

If an animated agent called 'James' is outgoing and unfriendly, it is formalized as

$$\textit{personality\_type(james,outgoing,unfriendly)}$$

Our conception of personality does not support quantification at the moment (as in [39]). The first dimension, extraversion, is considered in the conversational manager: outgoing agents try to take turn in a conversation whenever possible, whereas introverted agents only respond when offered to take turn. The second dimension, agreeableness, is reflected by the linguistic style the agent chooses when responding to other agents as vocal parameters set according to the vocal effects of basic emotions (see below).

C. Roles, Conventions, and Social Networks

A significant portion of human conversation takes place in a socio-organizational setting where participating agents have clearly defined *social roles*, such as sales person and customer, or instructor and student (Moulin [29]). We assume that agents always know which role they play in a certain social context, and are able to play different roles successively. *Conventional practices* are guidelines (or restrictions) about socially appropriate behavior in a particular social setting. We can distinguish two kinds of guiding restrictions. First, *behavioral constraints* concern responsibilities, rights, duties, prohibitions, and possibilities associated with a social role; second, *communicative conventions* function as a regulatory for the agent's choice of verbal (and non-verbal) expressions in a given social context (Lewis [24]). We will focus on the choice of verbal and non-verbal behavior (emotion expression), depending on the agent's social role and personality.

Formally, in social or organizational groups roles are ordered according to a *power scale*, which defines the social power of an agent's role over other roles [29]. We consider just two such relations for agents L_i and L_j: L_i's role R_{L_i} is higher ranked on the power scale than that of L_j, R_{L_j}, or L_i's role and L_j's role have the same rank. The power relations are stored as facts of the form

$$\textit{social\_power(Rel,P,}R_{L_i}\textit{,}R_{L_j}\textit{)}$$

where the relation *Rel* is either ">" or "=", and the argument P is a value between 0 and 3. By definition, if the relation is "=", then P is zero. Otherwise, if the value of P is 3, for instance, then the rank of L_i is much higher than the rank of L_j. The *social network* is specified by the social roles and associated power relations.

Following Walker *et al.* [42], we also consider the *social distance* between two agents (used in their framework to determine an appropriate linguistic style). The distance relations have format

$$\textit{social\_distance(D,}L_i\textit{,}L_j\textit{)}$$

where $D \in \{0, 1, 2, 3\}$. If agents know each other very well, the social distance can be set to zero. Note that the social distance can be high even if two agents have the same rank. When agents interact, they do not only exchange information but also establish and maintain social relationships. Hence it is important that agents avoid introducing disharmony into a conversation [29] or threaten other agents' public face [42]. Based on the values of social power $P = p(L_i, L_j)$ and social distance $D = d(L_j, L_i)$, the agent L_i computes the *threat* θ from L_j of expressing a certain emotion by using the following simple equation

$$\theta = p(L_i, L_j) + d(L_j, L_i)$$

and storing the resulting value in facts of the form "*threat(T,L_i,L_j)*" (see also [42]). High values for θ lead to the agent's suppression of the expression of its emotional state, whereas low values allow the agent to show its emotional state. Additionally, we explicitly condition responses on whether an agent L_i respects the conventions associated with its social role towards the role of L_j, by stating (for simplicity, agents are identified with their roles) "*respects_conventional_practices(L_i,L_j)*". If conventional practices are not respected, the following fact is added: "*violates_conventional_practices(L_i,L_j)*".

This is of course a very simple view of a social network but, as shown below, it already allows us to explain various phenomena in actual conversations. Further accuracy can be included in a straightforward fashion, e.g., by considering the *degree* to which conventional practices are respected or violated. In this way, social interaction can be made idiosyncratic as it would allow to define characters that have varying tendency to act according to their social role.

D. Emotion Expression

Emotional behavior can be conveyed through various channels, such as facial display (expression), speech and body movement. The so-called 'basic emotions' approach (Ekman [15]) distills those emotions that have distinctive (facial) expressions associated with them and seem to be universal: *fear*, *anger*, *sadness*, *happiness*, *disgust*, and *surprise*. More precisely, Ekman prefers to talk about (basic) emotion families. Thus it is consistent to have many expressions for the same basic emotion. It is of prime importance in our work to clearly distinguish between emotional states and emotion expression. In order to avoid confusion and since there is only a limited number of comprehensive 'emotion words', we use *slanted* when referring to basic emotions instead of *italics* for emotional states.

Characteristics of basic emotions include, e.g., quick onset (emotions begin quickly) and brief duration, which clearly distinguish them from other affective phenomena such as moods, personality traits, or attitudes. Note that only enjoyment (*happiness*) and possibly *surprise* are 'positive' emotions. The enjoyment family covers amusement, satisfaction, sensory pleasure, pride, thrill of excitement, and contentment. Interestingly, the positive emotions do not have a distinct physiological signal. Ekman explains this by referring to the minor relevance of positive emotions in evolution.

The 'basic emotions' approach provides a useful list of emotions as an emotion family inventory for animated agents as it explicitly relates emotion to behavior. Besides facial expressions, Ekman also suggests other signals, such as speech and body movement to express emotions. Most importantly, Murray and Arnott [31] describe the vocal effects of Ekman's five basic emotions. E.g., if a speaker expresses *happiness*, then his or her speech is typically faster (or slower) and higher-pitched, whereas a speaker expressing *sadness* usually uses slow and lower-pitched speech. When using those results for speech synthesis, however, we have to be aware that speech contains a variety of information, including cues to speaker identity, affect, and emphasis. Picard [35] reports on work saying that emotional arousal is communicated by pitch and loudness, whereas the realization of emotional valence needs more sophisticated cues, e.g., rhythm, and also non-acoustic cues such as facial expression.

V. Social Filter Programs

Basically, a social filter program consists of a set of rules that encode qualifying conditions for emotion expression. The program acts as a 'filter' between the agent's affective state and its rendering in a social context, such as a conversation. Hence, we prefer to talk about social *filter* programs rather than control programs (as in Gratch [18]). We consider the agent's personality and the agent's social role as the most important emotion expression qualifying conditions.

By way of example, let us explain the *angry-at* emotion type. Assume that a student is angry at his or her instructor because s/he is treated in an unfair way, which she considers as blameworthy. How will s/he react to his or her instructor? Presumably s/he will nod, showing that s/he understood the instructor's point of view, and try to argue that the instructor's interpretation is wrong, in a calm voice with a rather neutral facial expression.

The student's behavior—*suppressing* the expression of his or her emotional state—can be explained in at least two ways. *First*, s/he might have personality traits that characterize him or her as very friendly. *Second*, and probably more important in this scenario, s/he might be aware of his or her social role as a student which puts behavioral restrictions on his or her answer to the instructor.

A. Social Filter Rules

In the following, we will give some examples of social filter rules. We assume that emotion expression (e.g., facial display or linguistic style) is determined by personal experience, background knowledge, and cultural norms (Walker *et al.* [42]), as well as the 'organizational culture' (Moulin [29]). Consequently, human agents determine the values of the social variables 'social power' and 'social distance'. Our rules are consistent with Brown and Levinson's theory of social interaction, as reported in [42]. The filter rules should be understood as descriptive and not normative, i.e., we try to establish social role awareness as a control state in Allen's [1] sense (as a predictor for behavior) rather than to formulate how humans and agents should behave.

If the conversational partner has more social power and/or the social distance is large (i.e., the threat θ is high), the expression of 'negative' emotions is typically suppressed, resulting in *neutral* emotion expression (see Fig. 1, top). The first and second condition of the (top) rule concern the social context, the third condition accounts for the output of the affective reasoner, the emotional state. Here, the agent's emotion expression is assumed to be independent of its personality traits. If an agent communicates with an agent that imposes low threat, personality traits come into effect. If an agent's personality traits can be characterized as unfriendly, the agent may show its emotional state without violating conventional practices. An agent might even express happiness about something which—*the agent believes*—distresses another agent (see Fig. 1, bottom), or show its anger (see Fig. 2, top). Moreover, if the agent violates conventional practices, it might show its emotional state even if the social and/or power distance is high (see Fig. 2, bottom).

An interesting consequence of our approach is the following. Since we clearly distinguish between emotional state and expression of emotion, we may add another possibility of an agent's misinterpretation of other agents' behavior. *First*, an agent never has direct access to others mental states, it can only have (possibly false) beliefs about their mental (e.g., emotional) states. *Second*, our distinction allows that agents *cheat* in their behavior by expressing a misleading

Emotion expression *neutral*: an agent $L1$ displays emotion expression *neutral* towards agent $L2$ **if**

the threat for $L1$ from $L2$ is high, $\theta > 3$

and $L1$ respects conventional practices towards $L2$

and [$L1$ is *angry-at* $L2$

or $L1$ is in state *distress*

or $L1$ is in state *fear*].

Emotion expression *happiness*: an agent $L1$ displays *happiness* towards agent $L2$ **if**

the threat for $L1$ from $L2$ is low, $\theta \leq 1$

and $L1$ respects conventional practices towards $L2$

and $L1$ has unfriendly personality

and $L1$ is *gloating* over $L2$.

Fig. 1. Partial definitions for expressions *neutral* and *happiness*.

Emotion expression *anger*: an agent $L1$ displays *anger* towards agent $L2$ **if**

the threat for $L1$ from $L2$ is low, $\theta \leq 1$

and $L1$ respects conventional practices towards $L2$

and $L1$ has unfriendly personality

and $L1$ is *angry-at* $L2$

Emotion expression *anger*: an agent $L1$ displays *anger* towards agent $L2$ **if**

the threat for $L1$ from $L2$ is high, $\theta > 3$

and $L1$ violates conventional practices towards $L2$

and $L1$ has unfriendly personality

and $L1$ is *angry-at* $L2$.

Fig. 2. Partial definition for expression *anger*.

emotion. E.g., an agent may express the emotion *sadness*, pretending to be in a distressed emotional state, although it is in the emotional state *joy*.

B. Violations of Conventional Practices

Despite the fact that obedience to conventional practices is expected in real-world socio-organizational settings, violations of conventional practices occur, and in particular, they seem to be dramatically more interesting. What happens if a instructor requests something from his or her student and the student refuses to follow the order? Consider the following conversation fragment.

Instructor (to student): *I would appreciate if you try to solve this problem as a homework.*

Student (to instructor): *Sorry, I have more important things to do.*

Here, the student violates conventional practices by ignoring the instructor's indirectly formulated order. This situation typically triggers a negotiation process where the agent with higher social rank makes his or her request more explicit, or the even directly refers to his or her role and associated rights, e.g., the power to request tasks from subordinates (Moulin [29]). From an emotion expression point of view, the instructor agent will typically be in an *angry-at* (the student) emotional state and may express its anger emotion due to its higher rank on the power scale. The crucial belief in the instructor agent's mental model is

$$blameworthy(conventional\_practice\_violated, instructor)$$

i.e., instructors (typically) consider it as blameworthy if their higher rank is ignored by students. On the other hand, we might imagine an instructor with friendly personality traits who is distressed about the student's behavior, and expresses the emotion *sadness*. Another possibility is that the instructor shows *neutral* emotion expression, if instructor and student are good friends and—according to the very organizational culture—typical conventional practices are not applicable.

C. Social Feedback

Social filter programs certainly support high social accuracy of agents' responses. However, we currently have no principled way to introduce *social feedback* mechanisms to our model, i.e., values for social power and distance are kept fixed during the interaction. However, when humans establish and maintain social relationships, e.g., the value of their social distance changes. This phenomenon is not reflected in our conversation model. Agents have no memory about the social implications of their interactions with a human or other agents. We can imagine that if a user repeatedly interacts with an agent, the agent should not only remember the user but also show more familiarity. For the case of sales agents, Bickmore and Cassell [4] even suggest strategies to decrease social distance, by letting the agent use small talk.

Another interesting phenomenon of human-human communication are *reciprocal feedback loops* where, e.g., one agent's use of polite linguistic style results in another agent adapting its linguistic style and adjusting its value of social distance. Reciprocal feedback might also temporarily overrule the impact of personality traits and emotional state.

VI. Role-Playing with Socially Intelligent Agents

Our interactive environment for English conversation training for Japanese speakers is based on the idea that users (language students) would enjoy getting involved in a role-play with animated, socially intelligent agents, and thereby overcome their uneasiness to converse in a foreign language. Although Japanese speakers have a good knowledge of English grammar and vocabulary, they mostly lack opportunities to actually listen to and speak English. Inspired by the Virtual Theater project (Rousseau and Hayes-Roth [39]), we implemented an interactive drama that offers the role of a customer in a virtual coffee shop. Unlike the use of socially intelligent agents as assistants for educational problem solving oriented games (e.g., Conati and Klawe [10]), our agents' task is to keep the user interested and motivated to interact with them repeatedly.

A. Implementation

The programmable interface of the Microsoft agent package [27] is used to run a virtual coffee shop session in a web browser (Internet Explorer 5). This choice put some serious restrictions from the outset: the characters available for this package have only a limited number of behaviors ('animations'), confining the realization of various emotions as well as some features of embodied conversational behavior (Cassell and Thórisson [8]). However, our goal is believablity on the level of socially adequate emotional response rather than lifelikeness (in the sense of realistic behavior). The Microsoft Agent package provides controls to embed animated characters into a web page based JavaScript interface, and

Fig. 3. The *Comi!Chara* Interface.

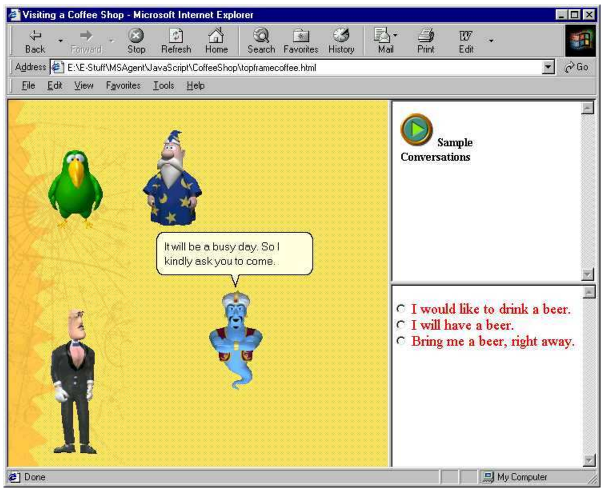

Fig. 4. Interaction involving user, James as waiter agent, Genie as manager agent, and two customer agents.

includes a voice recognizer and a text-to-speech engine. Prolog programs implement all reasoning related to conversation and environmental management and agents' mental models (affective and social reasoning). We use Jinni 2000 to communicate between Prolog code and the Java objects that control the agents through JavaScript code. Tarau's Jinni 2000 (Java Inference engine and Networked Interactor) can be used as a "scripting tool for gluing together knowledge processing components and Java objects in distributed applications" (BinNet Corp. [5]).

Our interactive environment for English conversation training is called *Comi!Chara*, which is conceived as a Japanese buzzword and abbreviates 'Communicate in English! With animated Characters' (see Fig. 3). The Comi!Chara tool is a Java applet which—in addition to standard GUI elements such as scroll-down menus—features an animated agent ('Charlie') guiding the user (language student) through the training environment. In particular, the user is motivated to start a role-playing session where he or she can interact with other characters. In a role-playing session, the user can promote the development of the conversation by uttering one of a set of predefined sentences that are displayed on the screen (see Fig. 4). Unlike the setup of Hayes-Roth's Virtual Theater project [39], the user does not have an avatar in the play. Animated characters will respond by synthetic speech, facial display, and gestures. However, there is no automatic synthesis of verbal or non-verbal behavior (as in Cassell [7], [4]). All of the agents' reactions are predefined and have to be encoded as functions in JavaScript. Depending on the output of the reasoning process related to an agent's mental model, the respective function is invoked and interpreted in the browser. The parameters for speech output are set in accordance with the vocal effects associated with the basic emotions (Ekman [15]). Facial display of characters is limited to the predefined 'animations' from the Agent package (e.g., 'pleased', 'sad'). To some extent, we also implemented conversational behavior (Cassell and Thórisson [8]). E.g., the animations 'confused' (lifting shoulders) and 'don't-recognize' (put hand to ear) are used if the user's utterance is not recognized.

B. Example Runs

We will illustrate our system by showing some example runs. In the first example, the user takes the role of a (friendly) customer who interacts with an unfriendly waiter agent James, who himself interacts with a friendly manager agent as an employee. Fig. 5 shows the annotated trace from our conversation system.

The communicative act of the customer has the form

com_act(customer,james,order_beer,polite,neutral,s0)

Since our agents do not understand English, a library is used to associate the user's utterance with an 'effect', e.g., that the regulations of the coffee shop are violated, and an evaluation of its linguistic style, such as polite, rude or neutral. Moreover, as an emotion (expression) recognition module is not part of our system, we set *neutral* as the default value for user input (but see the work of Picard [35]). The waiter's answer is formalized as

com_act(james,customer,refuse_beer,rude,anger,s1)

Similarly, a library is employed to generate the syntactic form of the agent's response. As described in Section 3, the environmental manager simulates the environment. In this example, it includes the fact

act(manager,appears,s2)

which triggers the waiter's reaction in situation $s3$. Hence, communicative acts can be performed both as a reaction to other communicative acts or events in the environment. In accordance with the beliefs in the waiter agent's mental model, the rules for affective reasoning, and the social filter program, the agent expresses its anger in situation $s1$. Here we assume that the customer's role is higher ranked than the waiter's role, the social power P has the value 1, and the social distance between waiter and customer is very high, D has value 3. Hence we obtain

threat(4,waiter,customer)

However, the waiter agent violates conventional practices and shows its anger. On the other hand, in situation $s5$, the waiter agent suppresses the expression of its emotional state, as the agent respects conventional practices towards the manager and the threat from the manager is high. The manager's role is higher than the waiter's role, the social power value is, say $P = 2$, and the social distance is high too, e.g., $D = 2$, so the value of θ is 4.

The second example run is a variation of the previous example where we assume an extrovert, friendly waiter who respects conventional practices towards customers but not towards his indifferent manager (see Fig. 6). There are limitless ways to vary social encounters between agents, even in the restricted context of a coffee shop environment. Consider, for instance, a situation where an unfriendly waiter shows rude behavior towards a customer, who turns out to be the waiter's new manager. A socially robust waiter agent will show a form of 'behavior switching', assuming that the agent respects conventional practices.

[s0] Customer: *I would like to drink a beer.* [User can also choose other beverages, and for each, he or she may select the linguistic style (polite, neutral, rude).]

[s1] James (to customer): *No way, this is a coffee shop.* [The waiter considers it as blameworthy to be asked for alcohol and shows his anger. He violates conventional practices towards the customer, as the social distance between waiter and customer is high.]

[s2] [The manager of the coffee shop appears.]

[s3] James (to manager): *Good afternoon, boss. May I take a day off tomorrow?* [Performs welcome gesture. Following conventional practices towards the manager, the waiter uses polite linguistic style.]

[s4] Manager: *It will be a busy day. So I kindly ask you to come.* [Uses polite linguistic style in accordance with his personality traits.]

[s5] Waiter: *Ok, I will be there.* [Considers it as blameworthy to be denied a vacation and is angry. However, he is aware of the threat from his boss and thus does not show his anger. Instead, he shows neutral emotion expression.]

Fig. 5. Conversation involving friendly customer, unfriendly waiter, and his friendly manager.

[s0] Waiter: *Welcome to our coffee shop! May I take your order?* [Starts the conversation because of his extrovert personality.]

[s1] Customer: *Bring me beer, right away.* [User chooses rude linguistic style.]

[s2] James (to customer): *I would really like to offer you a beer, believe me. Unfortunately, I am not allowed to serve alcoholic beverages here.* [Concludes that the customer is distressed and feels sorry for the customer.]

[s3] [The manager of the coffee shop appears.]

[s4] James (to manager): *Good to see you, boss. Tomorrow I want to take a day off.* [Performs welcome gesture.]

[s5] Manager: *Actually, I need you tomorrow. Thank you.* [Uses neutral linguistic style for his request.]

[s6] Waiter: *You don't have to thank me because I will not be here tomorrow.* [Waiter is angry as the manager refuses to give him a vacation. Since the waiter does not respect conventional practices towards the manager, he expresses his anger and refuses to obey the manager's order.]

Fig. 6. Conversation involving unfriendly customer, friendly waiter, and his indifferent manager.

C. User Feedback

We conducted a rather preliminary experiment on the impact of agents featuring social role awareness. As in the example runs of the previous section, users would play the role of a customer in a coffee shop and interact with an animated agent portraying a waiter. The waiter agent interacts with a manager agent, a fellow waiter agent or a customer agent. Although our general goal is to employ the animated agents approach to language conversation training, the focus here is to show that (i) users can recognize whether the agents behave according to their social roles, and (ii) the animated agents' responses are believable. Five users were asked to rate the appropriateness of the agents' responses. Furthermore, we asked them whether they think the agents' reactions could occur in real-world situations.

As we expected, users could identify the social roles played by the animated agents, which are easily detected in the coffee shop environment. Users could also recognize when conventional practices have been violated. However, answers varied when asked for what they think went wrong in case of violation: including, that the agent is in a bad mood, or does not like the boss (or customer agent). When we run the experiment with agents that only reason about emotion and personality (i.e., without social role awareness), users would generally not consider those agents as 'unbelievable', but they expected to get hints regarding the motivation for the violation of conventional practices, and appropriate reactions from the other agents. Our (early) findings reinforce the belief that an agent should show overall consistency in its behavior in order to come across as believable and that social role awareness facilitates consistency. Users generally enjoyed interacting with the animated agents, and were interested to experience the coffee shop scenario with different agent settings (personality traits, social parameters). They also felt compelled to figure out 'what kind of character' an animated agent might be.

The same users were also asked to interact with the *Native World* language training tool [32] (see Section 2). At the first encounter, users preferred the video clips with recorded humans to our animated agents. However, when asked to interact with *Native World* multiple times, users quickly lost their motivation to use the tool, for the obvious reason that they were presented the same video clip each time they gave the same input. Our animated agents, by contrast, would give a variety of different responses, depending on the contents of their mental model, and hence be more attractive as conversational partners.

VII. Discussion and Future Work

Our work aims to account for an important feature of human-human communication, namely social role awareness, that seems to have strong influence on our ways of emotion expression and our behavior in general. Social role awareness is approached from the viewpoint of the believability of animated characters as socially intelligent agents. It is shown that this feature of social interaction may explain phenomena such as suppressing (the expression) of emotions, as well as other forms of 'cheating' about an agent's emotion. As such, social role awareness can significantly contribute to the design of socially robust and dramatically interesting characters (as in Hayes-Roth *et al.* [21] or Rousseau and Hayes-Roth [39]). More recently, sensitivity to socio-organizational contexts is also pointed out as a crucial issue in military training simulations (Gratch and Hill [19], Gratch [18]). Here, interesting conflicts can arise between an agent's goal (or self-interest) and role-specific duties imposed by orders from a commander agent.

Although we believe that social role awareness makes animated agents more 'socially robust', our approach suffers from several shortcomings. In the following, we discuss work relevant to future refinements of our approach.

- **Social reasoning and planning.** Gratch [17] interleaves emotional reasoning with an explicit planning model. There are obvious advantages of considering an agent's plans, e.g., 'prospective' emotional states such as hope and fear assume reasoning about future situations and typically induce plan generation or the modification of current plans. Similarly, social reasoning would benefit from an explicit representation of plans. Often, an agent's choice of emotion expression depends on the state of its current plan: e.g., if an employee agent plans to get fired, its violation of conventional practices towards its manager should be seen in the light of this high-level goal, and not be considered as part of the agent's concept of its social role. Gratch's model of emotional reasoning, called *Émile* [17], provides other important desiderata for the design of animated agents, which should eventually be considered in our approach as well, e.g., intensity of emotion types and dynamics of emotion.

- **Social action.** Besides appropriate emotion expression, other (possibly more important) behavioral constraints apply to socio-organizational settings. The role of an agent is associated with certain responsibilities, rights, duties, prohibitions, and decision power (Moulin [29]). A broader perspective of social agency will require explicit representations of those behavioral constraints, as well as formalisations of social concepts such as 'commitment'. In this respect we may heavily draw on well-established research work on multi-agent systems and distributed artificial intelligence (e.g., Jennings [22], Castelfranchi [9]).
- **Social communication.** As mentioned throughout the paper, an obvious weakness of our approach is that we do not provide an explicit formalization of speech acts. Consequently, all of the dialogue contributions have to be carefully hand-crafted. In fact, we employed a simplified version of Moulin and Rousseau's conversation model (see also Moulin [29] and Bouzouba and Moulin [6]). In [29], Moulin introduces a new notation for speech acts that is tailored to communication in socio-organizational settings, in particular *conversational schemas* that allow an agent to select speech acts in accordance with communicative conventions. In addition, we should consider the *linguistic style (LS) strategies* discussed by Walker *et al.* [42]. Those strategies determine semantic content, syntactic form and acoustical realization of a speech act, qualified by the social situation. Application of LS strategies supports social interactions that allow agents to maintain public face (i.e., autonomy and approval). If speaker and hearer have equal social rank, 'direct' strategies can be applied (e.g., "Bring me a beer!"). On the other extreme, if the rank distance is very large, 'off record' strategies are chosen (e.g., "Someone has not brought me a beer.").
- **Social dynamics.** The most notable shortcoming of our approach—especially as a framework for socially intelligent agents—is that it does not adequately account for the social dynamics between agents (including human agents). As pointed out, e.g., by Dautenhahn [12] (see Introduction), a key aspect of agents' social intelligence concerns their ability to develop and maintain (social) relationships. In other words, there should be emergent phenomena resulting from agents' interactions. This is currently not the case in our approach.

VIII. Conclusion

In this paper, we propose to integrate social reasoning to mental models of animated agents, in addition to an affective reasoning component. The novel aspect of our work is that we explicate the social role of agents and associated constraints on emotion expression, which allows for enhanced believability of animated characters beyond reasoning about emotion and personality. We believe that considering the social dimension in animated agents approaches adds value for the following reasons:

- *Believability.* It may increase the illusion of life, which is often captured by reasoning about emotion and personality only.
- *Social Communication.* By respecting an important feature of human conversation, it adds 'social robustness' to agent-human and inter-agent communication.

- *Explanatory power.* It explains the frequent mismatch between the output of emotional reasoning (the emotional state) and emotion expression.

We have described web-based interactive drama scenarios featuring animated agents as an entertaining test-bed to experiment with new capabilities of agents. By considering the issues described in the discussion section, we hope to gain a better understanding of the social dimension in communication.

References

[1] Steve Allen. Control states and motivated agency. In E. André, editor, *Proceedings of the i3 Spring Days'99 Workshop on Behavior Planning for Life-like Characters and Avatars*, pages 43–61, 1999.

[2] Elisabeth André, Martin Klesen, Patrick Gebhard, Steve Allen, and Thomas Rist. Integrating models of personality and emotions into lifelike characters. In *Proceedings International Workshop on Affect in Interactions. Towards a New Generation of Interfaces*, 1999.

[3] Joseph Bates. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, 1994.

[4] Timothy Bickmore and Justine Cassell. How about this weather? social dialogue with embodied conversational agents. In K. Dautenhahn, editor, *Socially Intelligent Agents: The Human in the Loop (Papers from the 2000 AAAI Fall Symposium)*, pages 4–8, 2000.

[5] BinNet Corp. *Jinni 2000: A high performance Java based Prolog for agent scripting, client-server and internet programming*, 2000.

[6] Karim Bouzouba and Bernard Moulin. Connaissances implicites et sociales: Dialogisme des interactions discursives. In B. Moulin, S. Delisle, and B. Chaib-draa, editors, *Analyse et Simulation de conversations: De la théorie des actes de discours aux systèmes multiagents*, pages 203–242. L'Interdisciplinaire, 1999.

[7] Justine Cassell. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In J. Cassell, S. Prevost, J. Sullivan, and E. Churchill, editors, *Embodied Conversational Agents*. The MIT Press, 2000.

[8] Justine Cassell and Kristinn R. Thórisson. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13(4–5):519–538, 1999.

[9] Cristiano Castelfranchi. Modeling social action for AI agents. In *Proceedings 15th International Conference on Artificial Intelligence (IJCAI-97)*, pages 1567–1576, 1997. Invited Talk.

[10] Cristina Conati and Maria Klawe. Socially intelligent agents to improve the effectiveness of educational games. In K. Dautenhahn, editor, *Socially Intelligent Agents: The Human in the Loop (Papers from the 2000 AAAI Fall Symposium)*, pages 65–70, 2000.

[11] Kerstin Dautenhahn. The art of designing socially intelligent agents – science, fiction, and the human in the loop. *Applied Artificial Intelligence, Special Issue on Socially Intelligent Agents*, 12(7–8):573–617, 1998.

[12] Kerstin Dautenhahn. Embodiment and interaction in socially intelligent life-like agents. In C. L. Nehaniv, editor, *Computation for Metaphors, Analogy and Agent*, LNAI 1562, pages 102–142. Springer, 1999.

[13] Kerstin Dautenhahn. Socially intelligent agents and the primate social brain. towards a science of social minds. In K. Dautenhahn, editor, *Socially Intelligent Agents: The Human in the Loop (Papers from the 2000 AAAI Fall Symposium)*, pages 35–51, 2000.

[14] Kerstin Dautenhahn, editor. *Socially Intelligent Agents: The Human in the Loop (Papers from the 2000 AAAI Fall Symposium)*. The AAAI Press, Technical Report, 2000.

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

[16] Clark Elliott. *The Affective Reasoner. A process model of emotions in a multi-agent system.* PhD thesis, Northwestern University. Institute for the Learning Sciences, 1992.

[17] Jonathan Gratch. Émile: Marshalling passions in training and education. In *Proceedings 4th International Conference on Autonomous Agents (Agents'2000)*, pages 325–332, 2000.

[18] Jonathan Gratch. Socially situated planning. In K. Dautenhahn, editor, *Socially Intelligent Agents: The Human in the Loop (Papers from the 2000 AAAI Fall Symposium)*, pages 61–64, 2000.

[19] Jonathan Gratch and Randall W. Hill. Continuous planning and collaboration for command and control in joint synthetic battlespaces. In *Proceedings 8th Conference on Computer Generated Forces and Behavioral Representation*, 1999.

[20] Anthony Guye-Vuillème and Daniel Thalmann. Requirements for an architecture for believable social agents. In *Proceedings 4th International Conference on Autonomous Agents (Agents 2000)*, pages 48–49, 2000.

[21] Barbara Hayes-Roth, Robert van Gent, and Daniel Huber. Acting in character. In R. Trappl and P. Petta, editors, *Creating Personalities for Synthetic Actors: Towards Autonomous Personality*. Springer Verlag, 1997.

[22] Nick R. Jennings. Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3):223–250, 1993.

[23] James C. Lester, Sharolyn A. Converse, Brian A. Stone, and Susan E. Kahler. Animated agents and problem-solving effectiveness: A large-scale empirical evaluation. *Artificial Intelligence in Education*, pages 23–30, 1999.

[24] David Lewis. *Convention. A Philosophical Study*. Harvard University Press, 1969.

[25] Michael Mateas. An Oz-centric view of interactive drama and believable agents. Technical Report CMU-CS-97-156, School of Computer Science, Carnegie Mellon University, 1997.

[26] Helen McBreen, Paul Shade, Mervyn Jack, and Peter Wyard. Experimental assessment of the effectiveness of synthetic personae for multi-modal e-retail applications. In *Proceedings 4th International Conference on Autonomous Agents (Agents'2000)*, pages 39–45, 2000.

[27] Microsoft. *Developing for Microsoft Agent*. Microsoft Press, 1998.

[28] Dave Moffat. Personality parameters and programs. In R. Trappl and P. Petta, editors, *Creating Personalities for Synthetic Actors*, pages 120–165. Springer, 1997.

[29] Bernard Moulin. The social dimension of interactions in multiagent systems. In W. Wobcke, M. Pagnucco, and C. Zhang, editors, *Agents and Multi-Agent Systems, Formalisms, Methodologies and Applications*, pages 109–122. Springer Verlag, 1998.

[30] Bernard Moulin and Daniel Rousseau. An approach for modeling and simulating conversations. In D. Vanderveken and S. Kubo, editors, *Essays in Speech Act Theory*. John Benjamins, 2000.

[31] Iain R. Murray and John L. Arnott. Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication*, 16:369–390, 1995.

[32] Oki Software Co., Ltd., Native World Ver. 2.0, 1999.

[33] Paul O'Rorke and Andrew Ortony. Explaining emotions. *Cognitive Science*, 18:283–323, 1994.

[34] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[35] Rosalind W. Picard. *Affective Computing*. The MIT Press, 1997.

[36] Helmut Prendinger and Mitsuru Ishizuka. Carrying the role-playing metaphor to interactive learning environments. In *Proceedings International Conference on Intelligent User Interfaces (IUI-2001)*, 2001. Accepted as Poster/Demo. To appear.

[37] Byron Reeves and Clifford Nass. *The Media Equation. How People Treat Computers, Television and New Media Like Real People and Places*. CSLI Publications, Center for the Study of Language and Information. Cambridge University Press, 1998.

[38] Raymond Reiter. The frame problem in the situation calculus: a simple solution (sometimes) and a completeness result for goal regression. In Vladimir Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation. Papers in Honor of John McCarthy*, pages 359–380. Academic Press, Inc., Boston, New York, 1991.

[39] Daniel Rousseau and Barbara Hayes-Roth. A social-psychological model for synthetic actors. In *Proceedings 2nd International Conference on Autonomous Agents (Agents'98)*, pages 165–172, 1998.

[40] Stuart J. Russell and Peter Norvig. *Artificial Intelligence. A Modern Approach*. Prentice Hall, Inc., Upper Saddle River, New Jersey, 1995.

[41] Virtual Personalities Inc., Verbot Sylvie Ver. 3.04, 1998.

[42] Marilyn A. Walker, Janet E. Cahn, and Stephen J. Whittacker. Improvising linguistic style: Social and affective bases for agent personality. In *Proceedings First International Conference on Autonomous Agents (Agents'97)*, 1997.