

SOCIAL BONDING IN TALKING WITH SOCIAL AUTONOMOUS CREATURES

Michio Okada, Noriko Suzuki and Masaaki Date

ATR Media Integration & Communications Research Laboratories,
2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-0288 Japan
e-mail: {okada, noriko, date}@mic.atr.co.jp

ABSTRACT

Cyberpets and artificial autonomous creatures become popular applications of spoken dialogue systems as well as of artificial life and artificial intelligence research. One of next targets in the fields is to give them social skills so as to be social entities in human-inhabited environments. To behave as social entities, they first have to establish the social bonding with social others, and to coordinate social distance with them. The abilities are also what conventional spoken dialogue systems lack. This paper discusses the social aspects in interacting between human and autonomous creatures, such as dialogue coordination, social bonding and social understanding. One of the points is how to establish and maintain a dynamic, structural coupling with social others. We are focusing the relationship between entrusting behavior and its grounding in our everyday activities. And then, we discuss on a model of the structural coupling with dialogue participants based on the notion of the entrusting behavior and its grounding. Finally, we show an example of artificial autonomous creatures that is implemented in multiple situated agents for the model.

1. INTRODUCTION

Autonomous software agents that are able to engage in real-time spoken dialogues with the users are increasingly populating interactive media. Cyberpets, sentient characters and artificial autonomous creatures become popular application of spoken dialogue systems as well as that of artificial life research [1-5]. However, the majority of developed spoken dialogue systems do not adequately address the social aspects in interacting between human and artificial creatures. One reason for this is that research efforts in spoken dialogue systems have concentrated on dialogue management for crucial for navigating user behavior to achieve goals. In the task-oriented dialogue, for instance, the meaning and role of the messages exchanged are expected not to be ambiguous or vague. Inquiries in the interactions are expected to deal with commands like orders from a policeman or a president. The interaction style makes the relationship between user and system asymmetry in which the social distance of them separates.

In this paper, we discuss the social aspects to achieve a symmetric relationship and a close social distance in talking with artificial autonomous creatures such as cyberpets and sentient characters. To do so, we first focus on communication manners and primary motivation in our spontaneous chatting as everyday activities. It should be noted that the primary motivation of our everyday conversation is not always to convey messages to others, but rather to make sure their social bonding and to coordinate social distance during the interactions. From the perspective, we are rethinking social aspects for artificial autonomous creatures and social skills to establish the social bonding with human so as to be social entities in human-inhabited environments.

2. INTERACTION MANNERS IN EVERYDAY DIALOGUE

We are focusing on the construction of an interaction manner in everyday dialogue, which engages in social exchanges that maintain dialogue coordination and social bonding.

2.1 Interaction Style in Spoken Language Systems

In designing communication systems, we have intended to build them that can provide efficient and accurate information exchanges among users. Spoken dialogue systems have been also studied so as to achieve smart human-computer interfaces that are able to correctly interpret the meanings of other's utterance, and produce adequate responses. In the type of interactions, we tacitly assume that each utterance conveys complete meanings between dialogue participants where the speaker's meanings are stated in a one-sided way. As a result, the interaction style in a conversation has an imperative tone of voice like a command-based interaction with computers. We, researchers in spoken dialogue systems, maybe take a normal attitude toward the interaction style in designing interface systems. It should be noted here that the relationship between human and computers is asymmetry and separates. Is far from the style of familiar participants in everyday conversation.

Recently, to improve usability and friendliness in human-computer interaction, the studies of interface agents [3] are gaining more attention. However, the

relationship between humans and the interface agents remains asymmetry in which people feels a distance with the interface agents.

2.2 Interaction Style in Everyday Conversation

Roughly speaking, spoken language has evolved suitable for everyday human-human interaction in which the relationship of the participants is close and symmetry. People can choose communication strategies according to the social distance, such as literal or tacit communication. In the sense, the use of the spoken language in the restricted command-style interaction with computers is likely to cause a sense of incongruity to users.

In casual chatting with friends, the relationship is basically symmetry and the social distance is not so separate. And the primary motivation of our everyday casual chatting with friends is not always to convey new information to each other, but rather to manage a social distance and communication strategies with dialogue partner. We term the basic relationship with humans *social bonding*. The notion of the social bonding and attachment behaviors comes from mother-infant interactions and proto-communication in developmental psychology. It can be applied to the case of relationships between human and artificial agents such as interface agents and artificial autonomous creatures.

3. STRATEGIES FOR SOCIAL BONDING

In order to investigate a basic mechanism of social bonding, we discuss here the nature of spontaneous behaviors, which entrust their meaning and value to the interactive environment. We first illustrate the basic framework for our everyday activities from a viewpoint of ecological psychology [8].

3.1 Entrusting Behavior

While taking a walk, nobody thinks about the meaning and value of each step prior to every step. People seem to be aware of the meaning and the value of their behavior while interacting with the environment. From an ecological view of point, behaviors are navigated by the "information" that emerges from interactions between the action and its environment. In spontaneous behavior, since we are not able to see the emerging behavioral meaning, we try to entrust the meanings of the behaviors to the environment in order to find them with prospective awareness.

We call this basic strategy of spontaneous behaviors *entrusting behavior*. On the other hand, the role of environment that embodies the behavioral meaning is called *grounding*. Spontaneous behaviors are organized from these two processes: entrusting behavior and grounding.

3.2 Entrusting Behaviors in Social Interaction

Such a perspective on spontaneous behaviors can be extended to social interactions like everyday conversation. Assume that you say "Hello" to your colleague as a greeting in the morning. If your colleague went away in complete disregard to your greeting, your "Hello" would not perform the greeting as expected. The meaning and value of your utterance are supported by the response of your dialogue partner. When you spontaneously utter a sentence to someone, you have a prospective awareness of the other person's response and the emerging meaning of your utterance. However, you can not see the complete meaning prior to a response, so that your utterance is always produced as a speculation. Such speculative utterances are regarded as a kind of entrusting behavior that entrusts the meanings and values of the utterances to social others. Simultaneously, this entrusting behavior requires your dialogue partner to be in a *conspiracy* with you and support and share the emerging meanings during the exchange. The conspiracy relation can be regarded as an origin of the social bonding.

3.3 Responsibility to Respond

Human beings unconsciously feel a sort of responsibility to respond to another's approach. It is too hard to ignore our colleague's greeting as an intended attitude. However, artifacts such artificial autonomous creatures do not feel the responsibility to respond to utterances for grounding the meaning. In order to generate an analogous interaction for human-human conversation, the creatures have to work on maintaining a mutually regulated interaction.

3.4 Conversational Field as Structural Coupling

In general, skilled coordinated behaviors are not always navigated by well-prepared plans, instead they are regulated by routine activities with coordinated action-perception cycles structurally coupled with the environment. For instance, well-trained car drivers do not try to investigate the meaning of turning their steering wheel, rather they are aware of its meaning from relations between unintentional motions of the steering and changes in the outside appearance. The unintentional motions are a kind of entrusting behavior, and the changes in appearance in response to motions work as a grounding process for the meaning and the value of the entrusting behavior.

The conversational field between dialogue participants emerges from the strained relation between the prospective awareness and the emerging meaning of the entrusting behavior. In everyday conversation, its implicit, primary goal is to maintain the conversational field that is a kind of structural coupling between

participants. We call it *self-motivated dialogue* to contrast it with a goal-oriented dialogue. According to the dialogue situation and participants, we choose the communication manners. In casual chatting, the nature of a self-motivated dialogue becomes dominant, and people seem to understand each other's thought by using tacit communication through a structural coupling to maintain the conversational field.

4. THE MODEL AND ARCHITECTURE

The mechanism, called *emergent computation*, provides the field for constraint satisfaction in an unpredictable environment. Skilled, coordinated behavior can be navigated by the internal field that emerges from interactions among situated-agents. Maintaining a self-sustained field in interactions with humans can produce the entrusting and its grounding behavior.

The realization of the emergent computation is based on a spread activation network of behavior modules (= situated-agents) similar to that of Maes [7]. We are applying the architecture to represent the entrusting behavior and its grounding process in everyday dialogue. Figure 1 shows a schematic view of our model for social bonding. The overall architecture for a dynamic action selection consists of a set of behavior modules (situated-agents) and two types of internal contexts: intentional context and environmental context. These autonomous behaviors can be regarded as a kind of entrusting behavior that entrusts the opportunities of activity to its environment, which consists of other behavior modules. The grounding in the behavior module is achieved when all preconditions are satisfied. As a result, the sustained relations organized by multiple goals and motivations of every behavior module construct a field that navigates skillful coordinated behaviors.

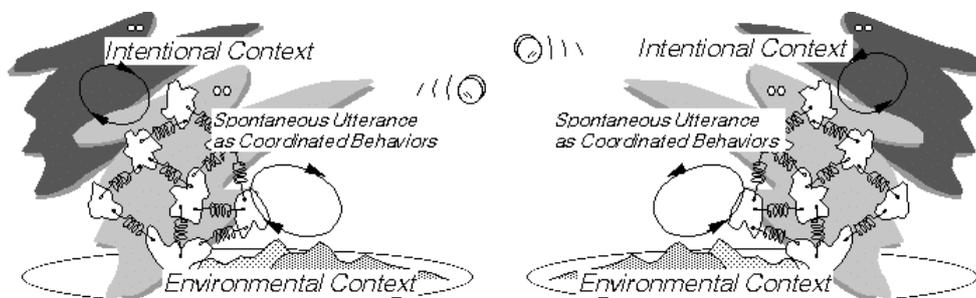


Figure 1. Schematic view of the model for social bonding.

4.3 Building Artificial Creatures

In order to investigate the basic mechanism and the nature of the social bonding discussed here, we are building several types of artificial creatures. Talking Eye [10] was built to investigate the nature of self-motivated dialogues that emerges from two communicating autonomous agents in virtual space (figure 2). Here, we are building a kind of embodied robotic agent in real world, because we believe that the agents require social effect of their actions to social others. And, a direct coupling between action and perception is important for grounding their actions with environments or others.

Figure 3 shows an overview of the artificial autonomous creatures we are developing. The appearance of them is like an infant with an eye, simple arms and legs that can appeal for people to interact with. These creatures communicate with each other, and communicate with a human through speech and vision. The body is made of an active CCD-camera, which form the central part of the creature. The motion of active camera can display eye's direction and the addressing of the creature, and provides social signals such as head turns and human nods.

4.4 Behavior Engine

The behavior engine of the creature consists of three sub-modules: perceptual system, behavior system, and motivational system. The perceptual system extracts vocal and visual events from the behaviors of social others. The vision system is implemented based on the behavior-based vision methodology using active cameras. This system is capable of detecting visual events such as nodding and expressions of agree or disagree from the movement and direction of the human face.

Speech recognition is performed using the continuous speech recognizer that has a vocabulary size of 300 words. By using syntactical templates, partial phrases are interpreted into content words and modality parts including sentence final particles and adjectives. Approximately 400 listed phrases are used for real-time responses for spoken utterance generation. These responses include various communicative acts such as requesting, informing, warning, suggesting, confirming, back-channel response, and agree/disagree. In the current implementation, the behavior system of the agents defines around 120 primitive behaviors both for

generating motions and utterances, and for regulating goals and motivations in the motivational module.

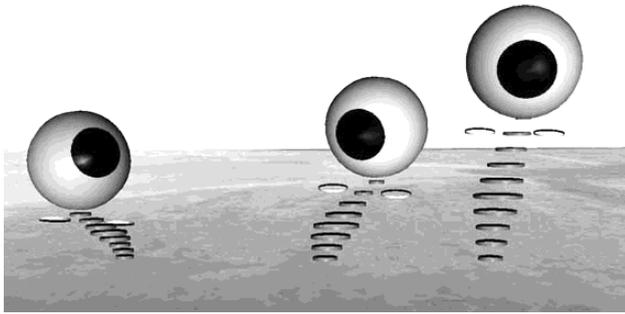


Figure 2. Overview of autonomous communicating creatures, Talking Eye.



Figure 3. Overview of artificial autonomous creatures in real world.

4.5 Driving Force for Interaction

The approach is demonstrated that is computationally simple and provides robust performance in human-computer interactions. Using this architecture, natural interaction with humans can be achieved without any explicit modeling of dialogue coordination. The driving force to yield the structural coupling between participants is likely that people is locally motivated to ground his/her own entrusting behavior through other's

grounding. As a result of their interaction, a dynamic, structural coupling between them emerges.

5. CONCLUSION

In our everyday conversation, the successive contributions of the participants are needed for maintaining their mutual relationship. So far, the efforts for joint activities are investigated as a basis for establishing mutual knowledge. We focused here that there are some strategies to coordinate their social distance in conversation. We discuss here the notion of entrusting behavior and its grounding that yield a social bonding with social others. The social bonding is not static, but rather dynamic, structural coupling that is built as mutual grounding process during continuous interactions such as everyday conversation. For artificial creatures, the motivation to establish structural coupling with social others will become a driving force for developing their communication skills.

6. REFERENCES

1. Bates, J. The role of emotion in believable agents, *Communications of the ACM*, 37(7), 1994.
2. Blumberg, B. and T. Galyean. Multi-level Direction of Autonomous Creatures for Real-Time Virtual Environments, *Proc. of SIGGRAPH 95*, 1995.
3. Maes, P. Agents that Reduce Work and Information Overload, *Communications of the ACM*, 37(7), 1994.
4. Cassell, J. et al. Modeling the Interaction between Speech and Gesture, *Proc. of the 16th Annual Conference of the Cognitive Science Society*, 1994.
5. Scassellati, B. Investigating Models of Social Development using a Humanoid Robot, *1998 AAAI Fall Symposium "Robots and Biology: Developing Connections"*, 1998.
6. Nass, C., Steuer, J. S., and Tauber, E. Computers are social actors. *Proc. of the CHI Conference*, 1994.
7. Maes, P. Situated Agents Can Have Goals, *Robotics and Autonomous Systems*, Vol.6 pp.49-70, 1990.
8. Reed, E. *Encountering the World, Toward an Ecological Psychology*, Oxford University Press, 1996.
9. Okada, M. *Hesitating Computer* (in Japanese), Kyoritsu Syuppan, 1995.
10. Suzuki, N., K. Ishii and M. Okada. Organizing Self-Motivated Dialogue with Autonomous Creatures, *Proc. of ICSLP-98*, 1998.

