Social Intelligence Design and Human Computing

Toyoaki Nishida

Graduate School of Informatics, Kyoto University
Sakyo-ku, Kyoto 606-8501, Japan
nishida@i.kyoto-u.ac.jp

Abstract. The central concern of Social Intelligence Design is the understanding and augmentation of social intelligence that might be attributed to both an individual and a group. Social Intelligence Design addresses understanding and augmentation of social intelligence resulting from bilateral interaction of intelligence attributed to an individual to coordinate her/his behavior with others in a society and that attributed to a collection of individuals to achieve goals as a whole and learn from experiences. Social intelligence can be addressed from multiple perspectives. In this chapter, I will focus on three aspects. First, I highlight interaction from the social discourse perspective in which social intelligence manifests in rapid interaction in a small group. Second, I look at the community media and social interaction in the large, where slow and massive interaction takes place in a large collection of people. Third, I survey work on social artifacts that embody social intelligence. Finally, I attempt to provide a structured view of the field.

Keywords: Social Intelligence Design, Conversational Informatics.

1 Introduction

The human is a social being. As an individual in society, a human should be able to manage relationships with other people and act wisely in a situation governed by an implicit or explicit set of social rules, in order to live a comfortable life. Such social skills range from an ability to communicate with and to learn from others to that of achieving goals by cooperating or negotiating with other people under a given social structure.

In the meanwhile, activities of humans constitute the human society. A collection of people may be considered to be intelligent if they can manage complexity and learn from experiences as a whole. The major concern of an organization is to maximize its problem solving power by coordinating the intellectual power of its members.

Social intelligence emerges from the synergy of individuals' social skill and the society's social structure. Both individuals' social skill and the society's social structure are indispensable to achieve social intelligence as a whole.

The advent of the network society has permitted people to interact with each other to affect social intelligence in a fashion unprecedented in history. On the one hand, it has brought about enormous benefits and conveniences. On the other hand, it has extended a dark side where a new technology is abused or disrupts human relations.

T.S. Huang et al. (Eds.): AI for Human Computing, LNAI 4451, pp. 190–214, 2007.

The central concern of Social Intelligence Design is the understanding and augmentation of social intelligence. Researchers with scientific or sociological concerns might be interested in uncovering the mechanisms of social intelligence as a function of individuals' social skill and the society's social structure. Those with engineering minds might want to design and devise advanced artifacts that will contribute to amplifying social intelligence.

Social Intelligence Design becomes central to human computing whenever one sheds light on situations in which a human is interacting with other people. Social Intelligence Design addresses a wide spectrum of phenomena concerning humans and their socially intellectual activities, ranging from establishing mutual intentions with nonverbal communications to new social media on the Internet.

In this chapter, I describe the conceptual framework of Social Intelligence Design, and present an overview of work on Social Intelligence Design presented at the Social Intelligence Design workshops.

2 Approaches to Social Intelligence Design

Social intelligence results from the bilateral interaction of intelligence attributed to an individual to coordinate her/his behavior with others in a society and that attributed to a collection of individuals to achieve goals as a whole and learn from experiences.

Social intelligence is ubiquitous in our daily life. Consider the following scenario.

> Scenario 1: K and J were coming to Kyoto for sightseeing. Although they found some nice famous places by Web search, they decided to go to a less famous, but quieter place with atmosphere that was recommended by local people. They posted their experience on the weblog.

Social Intelligence manifests in a variety of ways. It is evident that K and J relied on the experiences of other people on choosing the place for sightseeing. Recommendation and reputation were used as a primary means of knowledge resource to make a decision, in addition to catalog information about points of sightseeing spots. They might have preferred "local people" as a profile of the information source concerning the local affairs. They not only utilized the social information but they also contributed their own experience to the society. We also tend to rely on social heuristics such as the flow of people or a cluster of restaurants (such as the more restaurants, the more competitive and the nicer, or the longer history, the more reputation and the nicer).

In addition, there has been an enormous amount of tacit information "beneath the surface", such as believed to have happened at the nonverbal levels (gestures, facial expressions, voice tone, dressing, posture, etc), that are not explicit in the text, if you imagined what happened, say, before, during and after the conversation between the two groups. For example, there might have been a scene such as Figure 1 in the discourse. K and J might have been walking on the stone steps and looking at some interesting scene after J referred K to something to the right by speech and eye gaze; K might have realized and have started looking to the right; K and J might, then, have shared a sense of good communication by having realized that they had achieved a

Fig. 1. A scene of a fall day in Kyoto (painted by Mayumi Bono)

joint attention and start a conversation; nobody else might have paid much attention to them or might have ignored them out of courtesy; all might have kept quiet so not to disturb the quiet atmosphere.

Now, consider another scenario.

Scenario 2: T became a member of a new company. Soon after T came into the new office, he wanted to print his file. When he went to a printer room, he found a beautiful output was printed from a printer. He tried to set up parameters according to a manual, but he has not succeeded after many trials.

This scenario refers to an artifact (printer). Our modern life is surrounded by a huge number of artifacts of varying kinds. Although a naive thought about the introduction of artifacts is that they bring about great conveniences in our daily life, they may also result in a new stress. For example, we have to undertake various kinds of interaction with the artifact (e.g., setting parameters or issuing commands) in order to achieve a goal (e.g., having the content of a file printed on a paper), which may cause panic and loss of time, as illustrated in the above scenario.

In general, human-computer interaction induces social interactions. For example, the designer of the printer may embed the tips in addition to a normal usage manual. The tips may be implicit in the sense that it is implemented in a tacit way such as a shape or the light embedded in a button. The owner of a system may want to investigate customers' preferences by trying to infer usage patterns by computing statistics of the log data. In contrast, the user may communicate her/his message by actions such as choosing the product, answering customers' inquiries, or complaining. We can consider that artifacts are considered to be social communication media through which people exchange messages with each other or communicate with each other.

The social interaction features induced by an artifact may become more prominent as artifacts become more sensible and proactive to the user, autonomous to cope with novel situations, and capable of acting on behalf of the designer or owner. If this enterprise is successful, the artifacts may well complement the incompleteness of human's ability and knowledge, and any flaw of the current user support for complex machinery, as suggested in Scenario 2, may be dissolved. It will lead to the notion of artifacts as embodied social media.

Social Intelligence Design is a research field aimed at the understanding and augmentation of social intelligence. In what follows, I will overview work in Social Intelligence Design mainly published in the previous Social Intelligence Design workshops[1] from three perspectives suggested above. First, I will focus on the interaction from the social discourse perspective in which social intelligence manifests in rapid interaction in a small group. Second, I look at the community media and social interaction in the large, where slow and massive interaction takes place in a large group of people. Third, I survey work on social artifacts that embody social intelligence. Finally, I will attempt to provide a structured view of the field.

3 Interaction in Social Discourse

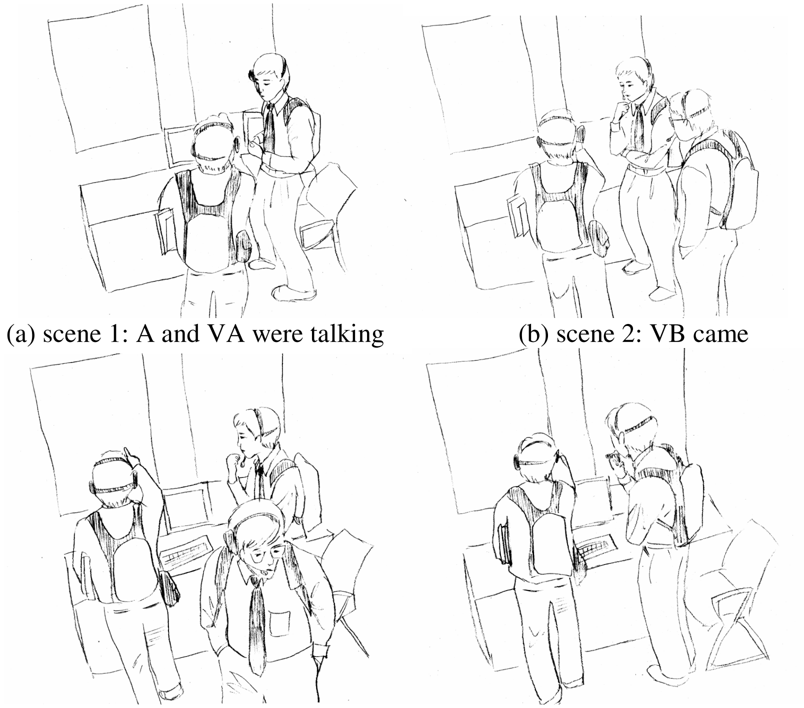
This perspective sheds light on structured interaction that takes place at the seconds-minutes order, at the conscious level. Collaboration protocol and story telling are key concepts at this level.

3.1 Social Interaction with Nonverbal Communication Means

People use nonverbal communication means in a proficient way to dynamically express their feelings, attitudes, intentions in addition to their thoughts. There is a tremendous amount of study on nonverbal communication behaviors including eye gaze, eyebrow, voice, facial expressions, head movement, hand gesture, body gesture, posture, distance, and so on. Kendon gave a structural description of gesture, most notably from the viewpoints of communication [Kendon 2004]. For example, he described in detail how greetings are made as a result of nonverbal interaction between social actors. McNeill studied relationships between gesture and thought [McNeill 05]. In addition to the development of measuring and coding schema for gestures, he presented a theory to characterize thought processes underlying gesture. He introduced a notion of growth point as "the smallest unit of the imagery-language dialectic", out of which a dynamic process of organization emerges. He characterized a growth point as "an empirically recoverable idea unit, inferred from speech-gesture synchrony and co-expressiveness." In addition, he introduced the notion of catchment to model multiple gestures with recurring form features to characterize the discourse segment to which the growth point belongs.

Some of these behaviors are intentionally produced by a communication partner for a communicative purpose, while others, such as a subtle correlation of eye gaze and mouth move, are not. In face to face communication, humans are considered to evaluate with each by sensing such unconscious, uncontrollable nonverbal behavior. There are several games such as poker where players are motivated to tell lies. Ueda and Ohmoto prototyped a real-time system that can discriminate lies by measuring gaze directions and facial feature points [Ohmoto 2006]. The system can measure

[1] http://www.ii.ist.i.kyoto-u.ac.jp/sid/



(a) scene 1: A and VA were talking (b) scene 2: VB came

(c) scene 3: VB left (d) scene 4: A and VA kept talking

Fig. 2. When the third person did not interrupt/join in an ongoing conversation (by courtesy of Dr. Mayumi Bono)

gaze directions and facial feature points, while allowing the user to move head position and orientation during the measurement and without requesting the user to place one or more markers on the face or preparing a face model in advance.

The discourse and task level information needs to be taken into account to see the role and function of each piece of communication behavior in the context of the entire discourse or task. Sacks introduced a concept of a turn-taking system in conversation, which suggests rules governing social interactions in conversation [Sacks 1974]. Multi-party face-to-face conversations such as those taking place at poster presentations are interesting as a subject of investigating mutual dependency between the task structure and the nonverbal communication behaviors. In the theory of participation status and framework, Goffman introduced such notions as side participants, bystanders, eavesdroppers in addition to the speaker and the addressee, in order to analyze the behavior of actors related to a conversation [Goffman 1981]. Bond used a framework of planning developed in Artificial Intelligence to model social behaviors as a set of joint steps with temporal and causal ordering constraints. An example of a joint behavior is grooming a pairwise affiliative behavior, which consists of four phases: orient, approach, prelude, and groom [Bond 2002].

Bono et al studied the dynamics of participation in face-to-face multi-party conversations [Bono 2004]. Figure 2 shows an example of a series of scenes at the poster session they investigated. In scene 1, an exhibitor (E) and a visitor (VA) were talking. During the talk, another visitor (VB) came in scene 2. Although VB came very close to E and VA in scene 3, VB left after around 27 seconds without exchanging conversation with A or VB. E and VA kept talking in scene 4.

It might be interesting to see the relationship between the utterances, distribution of eye gazes, and the arrangement of bodies. As for the distribution of eye gazes, E's eye direction was directed to VA, while VA was also gazing E in scene 1. Turn taking took place at an appropriate frequency. While one was making utterances as a speaker, he was looking at the hearer, making an explicit role assignment to the hearer. This interaction was maintained after VB arrived and VB was not chosen as a recipient of the utterance. As a result, VB had to stay as a side participant. This example scenario suggests that the role of conversational partner is determined proactively by the speaker.

Bono et al uncovered that the centrality of participation in a conversation is correlated with participants' interest in objects discussed as the topics of the conversation and that the more central a role one plays in a conversation, the more highly interested one is in the topic of conversation. In addition, they found that the frequency of interaction with exhibitors was a good indicator of visitors' interest in the corresponding posters, interest in the posters was proportionate with the time visitors spent being directly addressed by exhibitors rather than as side-participants, and addressee-hood of visitors manifested itself in the coordinated production of verbal backchannel responses.

Introduction of communication media may change the nature of communication. Mark and DeFlorio investigated how the use of wall-size, life-size HDTV was useful in overcoming the problems found in interaction with regular video-conferencing systems. The HDTV (High Definition TV) provides high-resolution images, conveyed to a far wider angle of the remote room with less distortion, than normal video-conferencing. Telephones were provided in each room, with telephone numbers, to support small group discussions (sidebar conversations) between the sites. They reported that people did not use exaggerated gestures or movements to convey expression through HDTV image in videoconferences. They also found that far fewer sidebar conversations across the geographically distant rooms were made during the videoconferences than during face-to-face conferences in the same room. They considered that the HDTV was mainly used for supporting the public conversations and as a result the team leader became a single primary channel of information for the group, as opposed to normal meetings, [Mark 2001]. In addition, they considered that the video image functioned as an awareness mechanism for activity in the other room, too.

3.2 Knowledge in Action

Gesture can be seen as external representations of abstract concepts which may be otherwise difficult to illustrate, as pointed out by Biswas and Fruchter [Biswas 2005]. They presented a framework for processing the captured video data to convert the tacit knowledge embedded in gestures into easily reusable and computable semantic representations. They prototyped a system called I-Gesture to investigate how people use gestures to express ideas. I-Gesture allows users to define a vocabulary of gestures for a specific domain, build a digital library of the gesture vocabulary, and mark up entire video streams based on the predefined vocabulary for future search and retrieval of digital content from the archive.

In order to capture and mine social communicative events for further knowledge reuse, Yin and Fruchter developed a prototype called I-Dialogue that captures the knowledge generated during informal communicative events through dialogue in the form of an unstructured digital design knowledge corpus. In addition, I-Dialogue adds structure to the unstructured digital knowledge corpus, and processes the corpus using an innovative notion disambiguation algorithm in support of knowledge retrieval [Yin 2007].

Skilled cooperative actions embody knowledge in co-action and can be seen as a form of social intelligence for sustainable interaction. Gill and Borchers considered people, tools, artifacts, and technologies to be dynamic representations of knowledge in joint design activities and aimed to study these in co-action [Gill 2003]. They observed and compared the behaviors of students to collaborate to design shared dorm living spaces using a normal whiteboard and two SMARTboards (electronic whiteboards that allowed for web browsing, typing up session notes, drawing mockups, and viewing their work). Body Moves was coded as the Parallel Coordinated Move (PCM). They obtained useful insights into the use of surfaces in collaboration. They reported that the ability to engage at the surface resulted in different strategies depending on the nature of the surface. For example, on the SMARTboard surface, body moves such as take-turn were used. As a result, the body field of the person acting was disturbed when the other one was entering it, and a reconfiguration of the engagement space was required. At the whiteboard surface, body moves such as attempt contact and focus were used to increase contact. In addition, further interesting phenomena were observed where participatory structure dynamically changed according to the problem solving status. They called it interactional dance.

3.3 Meeting Capture

Synthetic approaches at this level can be applied to the meeting room applications in order to better support meetings, such as real-time browsing, retrieval and summarization of meetings. Intelligent support of meetings can be characterized in the context of ambient intelligence defined as ubiquitous computing plus social and intelligent interfaces [Nijholt 2005b]. Issues in designing a meeting environment involve interpretation of events and activities in the environment, provision of real-time support, multimedia retrieval and reporting, autonomous and semi-autonomous embodied agents, controlling the environment, and its inhabitants. The M4 (multi-modal meeting manager) project was aimed at designing a meeting manager that can translate the information captured from microphones and cameras into annotated meeting minutes that allow for high-level retrieval questions, and for summarization and browsing. In order to track discussions, group actions such as presentations,

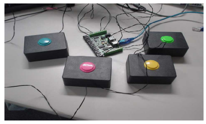


Fig. 3. Button-devices for human-assisted capture of conversation quanta [Saito 2005]

discussions, consensus and note-taking are modeled using Hidden Markov Models (HMMs). The actions of the individuals are recognized independently, and fused at a higher level. On the other hand, the augmented multi-party interaction (AMI) project aimed to develop new multimodal technologies to support human interaction in smart meeting rooms and remote meeting assistants. It aimed to enhance the value of multimodal meeting recordings and to make human interaction more effective in real time.

Saito et al reported a method of human-assisted conversation quanta acquisition using auxiliary button devices, as shown in Figure 3. In this conversation capture environment, each conversation participant is expected to press her/his button when s/he expects or finds a segment of conversation useful and hence intends to record the segment [Saito 2005]. The conversational situation itself is recorded during the conversation. Each participant can press the button to indicate the system her/his intention to mark. S/he can specify the in-point by quickly pressing the button n times to go back in the discourse $n \times d$ seconds, where d is a time unit for going back. Then, s/he keeps hold of the button until s/he thinks it is enough. The out-point will be set when s/he releases the button. As a result of an experimental investigation, five seconds was found to be an optimal value for d.

A meeting assistance environment called SKG Room as shown in Figure 4, was implemented. In this environment, a sphere surface of a spatial content aggregator called SKG is projected on a large screen. The users can have a conversation in front of the large screen. The conversation content can be captured by the human-assisted conversation contents capture environment. The captured conversation quanta are integrated on the SKG surface. In addition to the button devices, the touch screen was made available as a capture device, so that the user can touch the screen to signal her/his intentions.

In order to support the meeting in a more intelligent fashion, one might be interested in introducing pro-active meeting assistants that aim to assist the meeting process and thereby facilitate more effective and efficient meetings. Rienks et al discussed the requirement specifications of such meeting assistants and carried out a Wizard of Oz experiment [Rienks 2006]. A set of four different systems with varying intrusiveness levels was devised for the experiment. They were evaluated by two student committees of eight and seven members respectively. As a result, two versions that use voice samples together with a clock display when an item is due to be finished, something is off topic, a subject takes too long, or a discussion is unbalanced resulted in much more efficient meetings than those without pro-active meeting assistants, though less enjoyable.

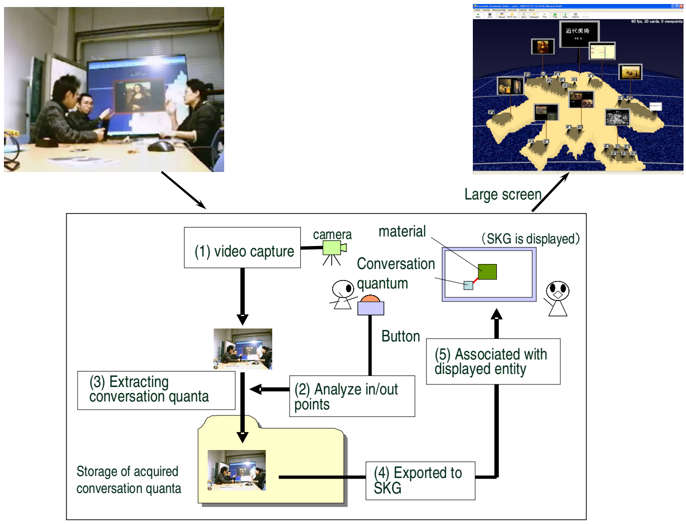


Fig. 4. SKG Room [Saito 2005]

3.4 Collaboration Support

Another key issue at this level is collaboration design. Fruchter proposes the analysis of collaboration support systems with three perspectives, physical spaces (e.g., "bricks"), electronic content ("bits"), and the way people communicate with each other ("interaction") [Fruchter 2001]. Fruchter introduced Brick & Bits & Interaction Hypothesis and Change Hypothesis and emphasize in designing their importance in collaboration technologies as follows:

Brick & Bits & Interaction Hypothesis: If we understand the relationship between bricks, bits, and interaction we will be able to

1. design spaces that better afford communicative events,
2. develop collaboration technologies based on natural idioms that best support the activities people perform, and
3. engage people in rich communicative experiences that enable them to immerse in their activity and forget about the technology that mediates the interaction.

Change Hypothesis: Any new information and collaboration technology will require change and rethinking of:

1. the design and location of spaces in which people work, learn, and play.
2. the content people create in terms of representation, media, interrelation among the different media, the content's evolution over time so that it provides context and sets it in a social communicative perspective.
3. the interactions among people in terms of the individual's behavior, interaction dynamics, new communication protocols, collaboration processes; relation between people and affordances of the space; and interactivity with the content.

The hypotheses were used in the Global Teamwork in Architecture, Engineering, Construction (A/E/C) offered at Stanford University.

4 Community Media and Social Interaction in the Large

This level is concerned with massive interaction in a society. It is slow and sometimes proceeds without much attention.

4.1 Understanding Community Media

There are several methodologies for analyzing macro interaction, involving social survey to collect information from the society, sociological methodology to interpret phenomena, social psychological methods, or statistic methods to derive a tendency. Miura and Shinohara investigated the communication congestion which is defined as the phenomenon of multiple topics simultaneously running when many participants are participating in the same chat session at the same time. [Miura 2004]. According to their study, high-density congestion can be regarded as leading to overloading of the participants' cognitive systems, making their remarks become less informative; moderate-density congestion with a relevant topic can be regarded as an activator of communication, particularly for experienced participants. Their research suggests that it might be possible to enhance the creativity of a group by adaptively changing the density of congestion.

A sociological approach may be applied to workplace design to facilitate communication and collaboration between mobile workers and their office base by ensuring compatibility between fixed and mobile, local and remote work areas. Rosenberg et al presented work on interaction space as the relationship between people, location, work task and process, based on the assumption that people jointly create an interaction space in which they work together towards achieving mutual understanding, and that the common ground is created in the interaction space that is shaped by spatial and organizational constraints as well as informational resources that determine the nature of the hybrid workplace. The investigation focused on the distance between people from three key perspectives [Rosenberg 2004].

In order to define a framework of standardized quantitative measurement of social intelligence, a notion of Social Intelligence Quantity (SIQ) was defined by combining the qualitative evaluation consisting of questionnaire and protocol analysis, and the quantitative evaluation consisting of the network log analysis, the factorial experiment and the standardized psychological scale [Yamashita 2002]. SIQ consists of SIQ-Personal and SIQ-Collective.

SIQ-Personal specifies the individual's personal attitudes to the society. SIQ-Personal is measured with the individual's information desire and intention to participate in the community. Matsumura investigated in detail the information desire and identified that the information acquisition desire consists of the interpersonal relation desire, the trend information acquisition desire, the information publication desire, the information monopoly desire, and the information acquisition desire [Matsumura 2004]. In order to investigate the structure of the individual's intention to participate in the community, Matsumura introduced seven factors (i.e., the intention

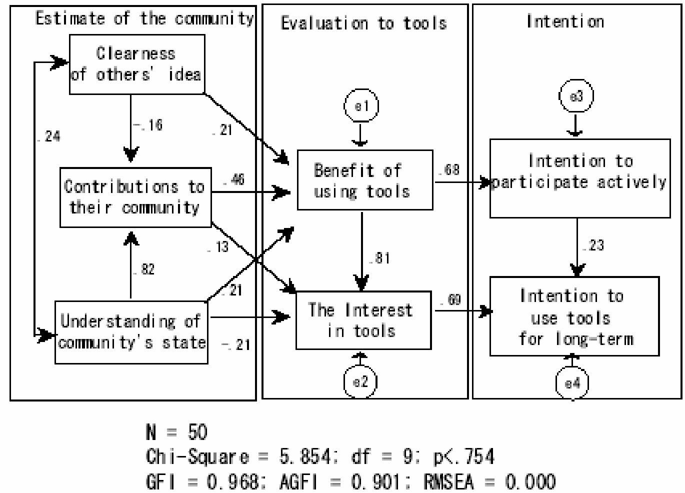


Fig. 5. A causal model concerning the intention of participating in a community and using a tool [Matsumura 2004]

of active participation, the intention of continuous participation, the benefits of using tools, the interest in the tool, the clarity of others' idea, the contribution to their community, and the understanding of the community's status), and proposed a causal model as shown in Figure 5. It indicates that the intention to participate in a community was influenced by the benefit received by using a communication tool. The benefit was affected by factors associated with understanding the state of a community. The individual's subjective contribution had a strong effect on the benefits of using a tool. The individual's subjective contribution depends on the understanding of the state of a community and the clarity of others' opinion. The model suggests that one should examine not only users' subjective evaluation of a communication tool but also their subjective evaluation of a community, in order to evaluate a communication tool for supporting a community.

SIQ-Collective represents the community's status of information and knowledge sharing or knowledge creation. Elaboration of SIQ-Collective is in progress. It is expected that SIQ-Collective may be defined by combining the amount of information in the community and the objective indices such as the diversity and the convergence of information.

4.2 Supporting Community

Digital Divide was addressed by Blake and Tucker. They discuss field trials that are underway in remote rural regions in South Africa in Tele-Health projects [Blake 2004]. There are a certain amount of synthetic techniques for producing awareness information by collecting information about what others are doing. FaintPop [Ohguro 2001] is a nonverbal communication device in which small icons of the user's colleagues are displayed. It allows the user to communicate her/his feeling towards her/his colleagues by using three types of touching. Fukuhara et al proposed a method called temporal sentiment analysis for finding sentiment of people on social problems according to the timeline. It provides a topic graph on a sentiment category specified by a user, and a sentiment graph on a topic specified by him or her. From these graphs, one can find socially concerned topics from the viewpoint of sentiment, understanding how people thought about social events.

IBM's Babble is a large scale discussion support system based on the idea of social translucence [Erickson 2007]. Erickson characterizes social intelligence as the ways in which groups of people manage to produce coherent behavior directed towards individual or collective ends. He introduced a device called the social proxy which is a minimalist graphical representation that portrays socially salient aspects of an online interaction, such as aspects of presence or other social activities of people. Based on experiences with implementing a number of design experiments, Erickson made six claims about designing with social proxies: everyone sees the same thing; no user customization; portray actions, not interpretation; social proxies should allow deception; support micro/macro readings; ambiguity is useful: suggest rather than inform; and Use a third-person point of view.

4.3 Knowledge Circulation in a Community

Circulating knowledge in a community is critical to the community not only because it enables knowledge sharing and formation of the common culture but also it increases the value of knowledge by going through an evolutionary process. In order to facilitate the knowledge circulation process in a community, we need to reduce overheads for knowledge circulation. We have developed a suite of technologies for conversational knowledge circulation. It consists of human-assisted acquisition of conversational contents, visual accumulation of conversational contents, and content-driven embodied conversational agents.

Conversation quantization was introduced as a conceptual framework underlying the entire package (Figure 6). The central idea in conversation quantization is to use a conversation quantum that encapsulates interaction-oriented and content-oriented views of a conversation scene. The interaction-oriented view sheds light on how each conversation partner coordinates her/his behaviors to communicate with the other. The content-oriented view, on the other hand, focuses on how meaning emerges from interaction.

Consider a conversational situation illustrated on the left-hand side of Figure 6, where one person on the left is explaining to another person on the right how to detach the front wheel from a bicycle. The interaction-oriented view of this scene is to see how participants interact with each other. It is quite probable that while the person on the left is explaining the operation with verbal and nonverbal modalities, the person on the right may also react to the explanation, say by looking or nodding.

A content-oriented view of this scene might be a proposition "in order to detach the front wheel from the bicycle, one must turn the lever to the right." Alternatively, it might simply be another proposition such as "A is talking to B," or "A is helpful." Thus, a content-oriented view may give a high-level, conceptual description of a conversation scene, while the interaction-oriented view grounds a content-oriented view. The interaction and content oriented views constitute a conversation quantum, as suggested on the right hand side of Figure 7.

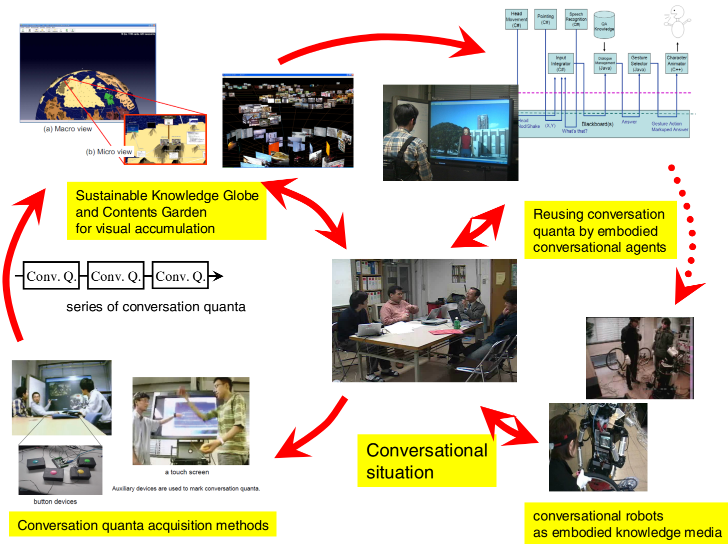


Fig. 6. Conversational Knowledge Circulation Enhancement Package

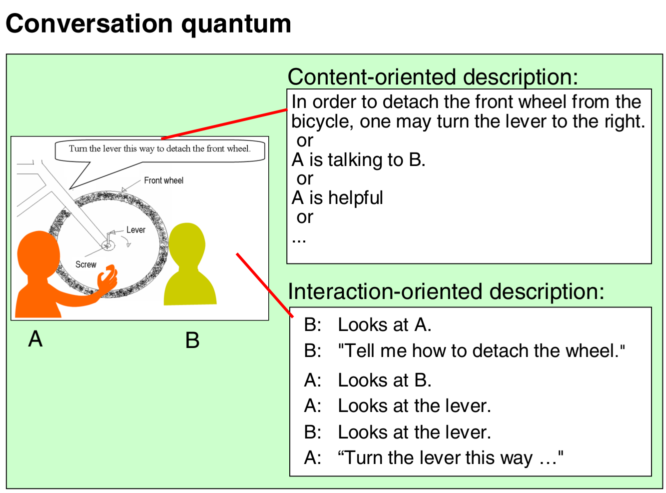


Fig. 7. A conversation quantum for a conversation scene

The scope of a conversation quantum is limited to a minimal conversation scene. This is because we would like to make a conversation quantum versatile so that it can be embedded in multiple conversation situations.

Capturing conversational contents from conversation scenes (materialization) involving identification of conversational scenes, recognition of communicative interactions and interpretation of communicative interactions is a challenging problem and hard to automate based on the current technology. A method of human-assisted conversation quanta acquisition using auxiliary button devices was developed. In order for a user to benefit from a large collection of conversation contents, a long-term relationship between the user and the conversation contents need to be established so that the conversation contents may coevolve with user's biological memory. Towards this end, we believe that supporting the spatial and temporal evolution of visually accumulated conversation contents ("knowledge landscape") is effective. Content-driven embodied conversational agents are used as a user interface. In the future, communicative robots will be used as a real world interface.

5 Social Artifacts

Insights obtained from Social Intelligence Design may be implemented as a social artifact such as an ECA or a conversational robot that can undertake social interaction with other social actors. (or -- that can interact socially)

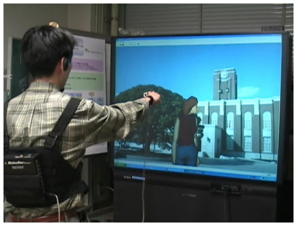
5.1 Embodied Conversational Agents

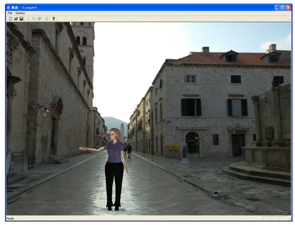
Embodied conversational agents (ECAs) are human-like characters that can interact with the user in a conversational fashion. As such, they are expected to induce a stronger form of the media equation [Reeves 1996] and to allow for natural social interaction with the user. In order to design an ECA that can mediate inter-human communication, we need to consider many aspects including ease of content management, a virtual environment that provide the context and referents for conversation, and implementation of higher order social concepts such as politeness, friendliness, humor, personality or trust.

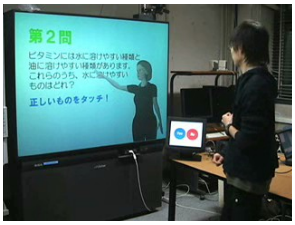
Nijholt proposed an information-rich virtual environment mimicking a real theater building where the user interacts with autonomous agents [Nijholt 2001]. The mental model of each autonomous agent is modeled in terms of beliefs, desires, plans, and emotions.

GECA (Generic ECA) is a platform (Figure 8) that can execute an immersive ECA system on multiple servers connected with each other by a computer network [Huang 2006]. GECA features:

- A general purpose platform (GECA platform) and a set of managing programs for mediating and transporting data stream and command messages between stand-alone ECA software modules that compose an ECA system interacting with human users in multiple modalities.
- A specification of a high-level protocol (GECAML) based on XML messages that are used in the communication between a standardized set of ECA components such as sensor inputs from the human users, inference engine, emotion model, personality model, dialogue manager, face and body animation, etc.
- An application programming interface (API) available on main-stream operating systems and programming languages for easily adapting ECA software modules to hook to the platform.


(a) a campus guide agent


(b) a tour guide agent



(c) a quiz agent

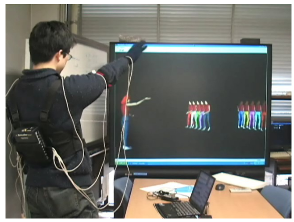

(d) a pedagogical agent

Fig. 8. ECAs built on top of GECA [Huang 2006]

The blackboard model is employed as the backbone to realize the weak inter-connectivity of the components, enabling the online switching of components and offline upgrading and maintaining of components. It allows logically isolated multiple blackboards to distribute information traffic concentrated on a logically single blackboard.

GECA has been implemented and applied for various applications involving a campus guide agent, a tour guide agent, a quiz agent, and a pedagogical agent for teaching cross cultural communication (Figure 8).

5.2 Communicative Robots

Understanding the functional and emotional aspects of nonverbal interaction is a key issue. Not all nonverbal behaviors are consciously controlled by the actors. By measuring the behavior of actors, we can find interesting aspects of social intelligence. Xu et al. built a human-human WOZ (Wizard-of-Oz) experiment setting that allows for observing and understanding how the mutual adaptation procedure occurs between human beings in nonverbal communication [Xu 2007]. The motivation of the WOZ experiment was to observe the nature of mutual adaptation of human-human interaction, which was intended to be implemented on robots. Figure 9 shows the experiment setting. One human subject (the actor) was asked to move around in a simulated mine field to clean target objects according to the instruction of another human subject (the instructor). The instructor was given the entire task map with the positions of all the obstacles on the map, including bombs, signals, targets, as well as the goal and orbit. The instructor would also need to perceive the current, exact position of the actor. The reward, in the form of sound effects (clearing target objects, exploding bombs/signals, reaching the goal, etc.) and scores were necessary for both the instructor and the actor. The mask and sunglasses were used to prevent

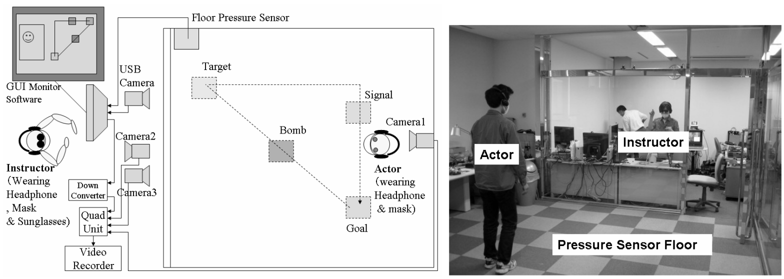


Fig. 9. A WOZ experiment setting for observing mutual adaptation [Xu 2007]

interference from other communication channels. The decision was made to prevent the subjects from communicating with eye gaze or facial expressions that were expected to be difficult to detect automatically when the observed behaviors would be implemented on robots. The instructor observes the movement of the actor by viewing the floor sensor information displayed on the computer screen. In order to enable the instructor to observe the actor's movements more clearly, a USB web camera was installed. As a result, three interesting findings were observed frequently. The first was called alignment-based action in which two participants align their actions while interacting with each other. The second was symbol-emergent learning in which the instructor used gestures such as symbol-like instructions when he interacted with the actor. The third was environmental learning in which communication between the participants became smoother in the subsequent trials and both participants adapted and improved their efficiency of movement.

5.3 Establishing Mutual Intention

Social intelligence at this level might be implemented as cohabiting with people. In order for a robot to cohabit with people, the robot should be able to be sensitive to subtle nonverbal behaviors of people. For example, in order for a robot to speak to a person who is making conversation with another person, it should follow a normal protocol people share in the society, that is, coming in the sight of the person, glance at her/his eye, wait for a call, and start talking as soon as it is called, just as is observed in inter-human communication [Kendon 1973]. The behavior should be modulated depending on how urgent the message is (Figure 10).

The key observation here is that the mutual intention formation between the client and the robots is established as a result of quick interactions using quick nonverbal interactions. The coordination of their behaviors is made by choosing options at each stage. If the gentleman does not want to be interrupted for a while, or the robot is called by some other urgent requests, they will act differently at respective stages, resulting in a different outcome.

We consider it is feasible to implement artifacts that can communicate with people with nonverbal communication means. The ability of forming and sustaining intention shared by participants (joint intention) is considered to be a primary goal in

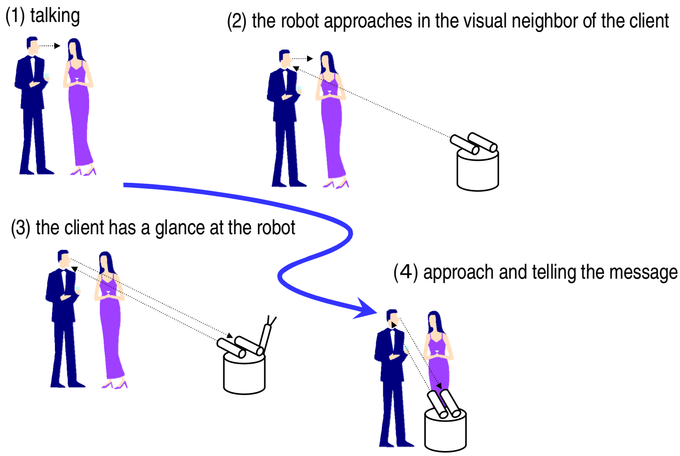


Fig. 10. How conversational robots should behave to speak to a person in conversation

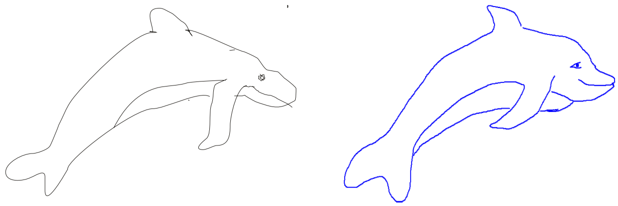
nonverbal communication. The communication schema we have in mind allows two or more participants to repeat observations and reactions at varying speeds to form and maintain joint intentions to coordinate behavior, which may be called a "coordination search loop" [Nishida 2006].

A study on embodiment communication with a mobile chair robot (a locomotive chair that can dynamically produce a means of allowing a person to get a place to sit down) suggests that although the users were all able to sit down on the chair as a result of coordinating behaviors, some users pointed out that the autonomous mobile chair should have communicated its intentions more explicitly [Terada 2001].

Listener and presenter robots were built by taking into account the lessons learnt from the mobile chair robot (Figure 11) [Nishida 2006, Ohya 2006]. The motivation underlying this research is to build a robot that can mediate knowledge among people. The listener robot interacts with an instructor to acquire knowledge by videotaping important scenes of her/his activities (e.g., assembling/disassembling a machine). The presenter robot, equipped with a small display, will then interact with a novice to show

(a) The listener robot (b) The presenter robot

Fig. 11. Robots as embodied knowledge media



(a) Initial drawing (b) Final drawing

Fig. 12. Is this what you intended to draw? [Mohammad 2007a]

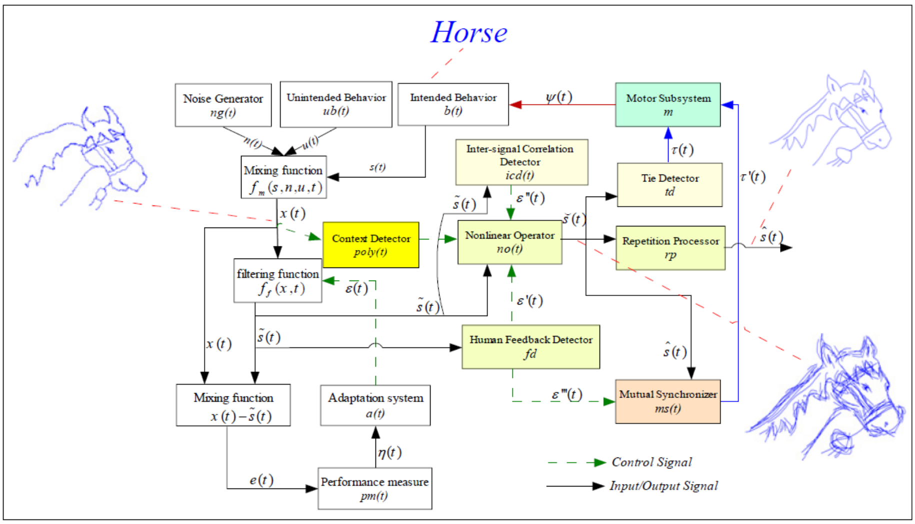


Fig. 13. The architecture of NaturalDraw [Mohammad 2007b]

the appropriate video clip in appropriate situations where this knowledge is considered to be needed during her/his work (e.g., trying to assemble/disassemble a machine).

From time to time, intention may not be clear, but it may become gradually concrete through interaction. A good example is drawing. Figure 12a and b were drawn by the same person using drawing software called NaturalDraw [Mohammad 2007a]. If the person was asked if the drawing in Figure 12a was what was wanted to draw, the answer would be no, for it was revised as in Figure 12b.

The idea of "coordination search loop" between the human user and the drawing software also helps here to derive the final drawing as a mutual intention between the user and the software [Mohammad 2007a]. Mohammad and Nishida used signal processing techniques based on the idea that intention can be best modeled not as a fixed unknown value, but as a dynamically evolving function. Interaction between two agents couples their intention functions creating a single system that co-evolves as the interaction goes toward a mutual intention state. They distinguish the intended signal from noise and unintended signal, and how the intended signal can be amplified through the repeated cycles of noise reduction and intention detection including automatic repetition detection that enables the system to interactively detect the existence of repetition in the drawing, as shown in Figure 13. Roughly, it is assumed that the input to the system is a mixture of intended behavior, unintended behavior, and noise. The noise is filtered out by repeatedly applying filtering functions. Low pass filters are not used, for they may remove intended high frequency signals. In the meanwhile, an inter-signal correlation detector and a human feedback detector are used to extract intended behaviors such as repetition and the result is shown to the user, who is expected to strengthen or weaken the system's interpretation by the feedback.

6 A Structured View of Social Intelligence Design

In this section, I attempt to encompass the entire business of Social Intelligence Design in a more structured fashion. First, I give a historical overview of the field. Then, I introduce several viewpoints to classify approaches to Social Intelligence Design.

6.1 Historical Overview of Social Intelligence Design

The ideas of social intelligence design developed along with a series of international workshops starting in 2001.[2]

The first workshop, SID2001, was intended to launch a new interdisciplinary study focusing on social intelligence, aimed at integrating understanding and development of new information and communication technology that may even result in the emergence of a new language and lifestyle [Nishida 2001]. The dual views of social intelligence, an individual's ability to be able to "manage relationship with other agents and act wisely in a situation governed by an implicit or explicit set of shared rules, based on an ability of monitoring and understanding of other agents' mental state" and an ability of a group of people to manage complexity and learn from experiences as a function of the design of social structure, were identified and equally emphasized.

The scope of Social Intelligence Design identified at SID2001 is summarized in Table 1. The primary concern of Social Intelligence Design was identified as design and implementation of novel communication means for mediating interaction among people and agents. The scope ranges from preliminary and preparatory interactions among people such as knowing who's who, to more intimate interaction such as collaboration. Supporting a group formation, collaboration, negotiation, public discussion or social learning is considered to be an important application of Social Intelligence Design. Theoretical aspects involve designing, deploying, and evaluating social intelligence support tools.

The second workshop, SID2003, focused on the theories to enhance the understanding and conceptualization of human cognition and interpersonal interaction to emphasize the practical aspects of Social Intelligence Design [Rosenberg 2005]. The major outcomes of the second workshop include: experimental approach to the

---

[2] http://www.ii.ist.i.kyoto-u.ac.jp/sid/

design of interactive workspaces, ethnographic studies of the introduction of technologies, the building of stories, methodologies for studying construction projects, the studies of online communities where mediated communication is a key vehicle for creating and maintaining social contact, building an interactive system that aids human decision and action, an analysis of the dynamism of the virtual community, and analysis of the effect of anonymous communication in online community.

The third workshop, SID2004, distinguished the following four themes [Nijholt 2006a]:

1. Natural Interactions - covering theory, modelling and analytical frameworks that have been developed with Social Intelligence Design in mind, including situated computation, embodied conversational agents, sociable artefacts, socially intelligent robots.
2. Communities - covering community media, communication patterns in online communities, knowledge-creating, network and anonymous communities.
3. Collaboration Technologies and tools - covering innovations to support interactions within communities, covering a range from knowledge sharing systems, multi-agent systems and interactive systems.
4. Application Domains - including design, workspaces, education, e-commerce, entertainment, digital democracy, digital cities, policy and business.

The major outcomes of the workshop involve the analysis of the effect of employing embodied conversational agents or communicative robots in human-artifact communication, analysis of presence in distributed workspaces, analysis of the effect of long-term use of distributed workspaces, analysis and modeling of meetings, analysis of network search activities, and an effort to compensate for the digital divide.

The fourth workshop, SID2005, added

- Multidisciplinary perspectives – exploring Social Intelligence Design at the intersection of different disciplines, such as, people-place-process, place-technology-interaction, that brings technology, work spaces, social behaviors and process aspects together.

to the fourth workshop [Fruchter 2007]. The major results involve a theory of improvisational social acts and communication, extending social intelligence design issues in extreme emergency and regular meeting room cases, analysis of the weblog community, supporting knowledge capture, sharing, transfer and reuse in collaborative teams.

The fifth workshop, SID2006, emphasizes the following three subjects:

1. Development, operation, and evaluation of support systems or tools for SID
2. Observation and modeling of psychological and behavioral processes of e-community
3. Social intelligence design by pilot program and computer-aided simulation.[3]

[3] http://www.team1mile.com/asarin/sid2006/

Table 1. The Scope of Social Intelligence Design identified at SID2001 [Nishida 2001]

1. Theoretical aspects of social intelligence design
   (a) Understanding group dynamics of knowledge creation
   (b) Understanding consensus formation process
   (c) Theory of common ground in language use
   (d) Attachment-based learning for social learning
2. Methods of establishing the social context
   (a) Awareness of connectedness
   (b) Circulating personal views
   (c) Sharing stories
3. Embodied conversational agents and social intelligence
   (a) Knowledge exchange by virtualized egos
   (b) Conversational agents for mediating discussions
   (c) A virtual world habited by autonomous conversational agents
   (d) Social learning with a conversational interface
   (e) Conversations as a principle of designing complex systems
   (f) Artifacts capable of making embodied communication
4. Collaboration design
   (a) Integrating the physical space, electronic content, and interaction
   (b) Using multi agent system to help people in a complex situation
   (c) Evaluating communication infrastructure in terms of collaboration support
5. Public discourse
   (a) Visualization
   (b) Social awareness support
   (c) Integrating Surveys, Delphis and Mediation for democratic participation
   (d) Evaluation of social intelligence
   (e) Network analysis
   (f) Hybrid method

The sixth workshop, SID2007, focuses on

1. Development, operation, and evaluation of support systems or tools for SID, which includes support systems and tools both for mediated remote interaction and support for face-to-face interaction;
2. Observation and modeling of psychological and behavioral processes with the aim of obtaining computational models of behavior and interaction;
3. Social intelligence implemented in interfaces, embodied agents, storytelling environments, (serious) gaming and simulation.[4]

6.2 Viewpoints to Classify Approaches to Social Intelligence Design

We might be able to classify approaches to Social Intelligence Design depending on which aspects of social intelligence are addressed.

(1) Micro in the small versus macro in the large

Social intelligence may manifest in a quick interaction among a small number of participants, as discussed concerning the first scenario. In contrast, it may be

[4] http://hmi.ewi.utwente.nl/sid07

observed in a slow interaction among a large number of participants. Typical examples are exchange of reputation and sharing of footprints on the Internet.

(2) Inter-human communication versus human-computer interaction

Social intelligence may be discussed in the context of inter-human communication. Computer-mediated communication falls into this category, where inter-human communication is mediated by ICT. In contrast, it may be discussed in the context of human-computer interaction, where a service reflects the intentions of various people including the owner, service provider, computer vender, and so on.

(3) Physical versus virtual

Social interaction takes place at the physical face-to-face interaction settings or in a virtual environment on the net. The phenomena may be more subtle and vague in the former case, for it is not clear exactly what features affect social intelligence.

(4) Speech acts

One might be interested in looking at the phenomena by looking at the speech acts involved in social interaction. Among others, cooperation, competition, and negotiation might be interesting. Higher level concepts include politeness or friendliness.

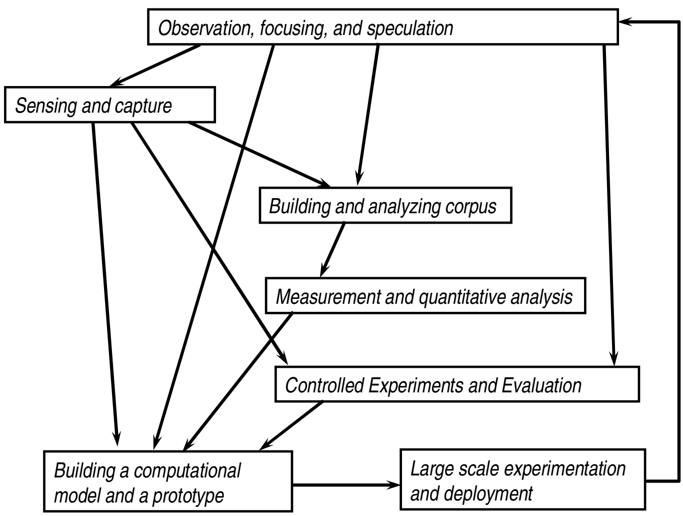


Fig. 14. Methodological Aspects of Social Intelligence Design

On the other hand, approaches to social intelligence design range from analysis motivated by scientific interests to synthesis motivated by engineering goals. Some might place much emphasis on understanding based on observation and investigation, trying to arguably identify and uncover critical aspects of social intelligence. Some others might be interested in modeling some aspects of social intelligence by following psychological disciplines. Yet others might be interested in building a system that may augment collective intelligence of a group or a community. It should be noted, however, that an integrated approach is very important in Social Intelligence Design. Analysis and synthesis should be closely integrated. Analysis is needed to predict and evaluate applications. Synthetic approach is needed to focus analysis. Individual researches in Social Intelligence Design might be located and interrelated in a big picture as shown in Figure 14, depending on which aspects are emphasized as primary focus.

7 Concluding Remarks

In this chapter, I described Social Intelligence Design as an endeavor of understanding and augmentation of social intelligence that might be attributed to both an individual and a community. Social Intelligence Design involves understanding and augmentation of social intelligence, which is defined as an association of intelligence attributed to an individual to coordinate her/his behavior with others in a society and that attributed to a collection of individuals to achieve goals as a whole and learn from experiences. In this chapter, I first surveyed work on interaction in social discourse perspective in which social intelligence manifests in rapid interaction in a small group. Then, I looked at the community media and social interaction in the large, where slow and massive interaction takes place in a large group of people. Then, I surveyed work on social artifacts that embody social intelligence. Finally, I attempted to provide a structured view of the field.

References

[Azechi 2001] Azechi, S. Community based Social Intelligence, SID-2001, 2001.

[Biswas 2007] Biswas, P. and Fruchter, R. Using Gestures to Convey Internal Mental Models and Index Mutimedia Contents, AI & Society, to appear.

[Blake 2007] Blake E. and Tucker, W. User Interfaces for Communication across the digital divide, AI & Society, to appear.

[Bond 2002] Bond, A. H. Modeling Social Relationship -- An agent architecture for voluntary mutual control, in: Dautenhahn, K. Bond, A. H., Canamero, L. and Edmonds, B. (eds) Socially Intelligent Agents -- Creating Relationships with Computers and Robots, Kluwer Academic Publishers, pp. 29-36, 2002.

[Bono 2004] Bono, M., Suzuki, N., and Katagiri, Y. An Analysis of Participation Structures in Multi-Party Conversations: Do interaction behaviors give clues to know your interest?, Journal of Japanese Cognitive Science Society, No.11, Vol. 3, pp. 214-227, 2004.

[Erickson 2007] Erickson, T. 'Social' Systems: Designing Digital Systems that Support Social Intelligence. To appear in AI and Society.

[Fruchter 2001] Fruchter, R. Bricks & Bits & Interaction, Yukio Ohsawa, Shusaku Tsumoto, and Takashi Washio (eds): Exploring New Frontiers on Artificial Intelligence - Selected Papers from the First International Workshops of Japanese Society of Artificial Intelligence -, Lecture Notes on Artificial Intelligence LNAI2253, Springer Verlag, December 2001.

[Fruchter 2005] Fruchter, R., Nishida, T., Rosenberg, D. Understanding mediated communication: the social intelligence design (SID) approach, AI & Society, Volume 19, Number 1, pp. 1-7, 2005.

[Fruchter 2007] Fruchter, R., Nishida, T., Rosenberg, D. Mediated Communication in Action: a Social Intelligence Design Approach, AI & Society, to appear.

[Gill 2003] Gill, S. P. and Borchers, J. Knowledge in Co-Action: Social Intelligence in Using Surfaces for Collaborative Design Tasks, presented at SID-2003.

[Goffman 1981] Forms of talk. Philadelphia: University of Pennsylvania Press, 1981

[Huang 2006] Huang, H., Masuda, T., Cerekovic, A., Tarasenko, K., Pandzic, I., Nakano, Y., and Nishida, T., Toward a Universal Platform for Integrating Embodied Conversational Agent Components, in: B. Gabrys, R.J. Howlett, and L.C. Jain (Eds.): KES 2006, Part II, LNAI 4252, pp. 220 – 226, 2006.

[Kendon 1973] Kendon, A. and Ferber, A. A Description of Some Human Greetings, In: Michael, R. P. and Crook, J. H. (eds.) Comparative Ecology and Behaviour of Primates, Academic Press, 1973.

[Kendon 2004] Kendon, A. Gesture: Visible Action as Utterance. Cambridge: Cambridge University Press, 2004.

[Mark 2001] Mark, G. and DeFlorio, P. HDTV: a challenge to traditional video conferences? Publish-only paper, SID-2001, 2001

[Matsumura 2004] Matsumura, K. The Measures for the Evaluation of Communication Tools: the Causality between the Intention and Users' Subjective Estimation of Community, in Proceedings of the 3rd Workshop on Social Intelligence Design (SID 2004), pp. 85-90, 2004.

[McNeill 05] McNeill, D. Gesture and Thought, Chicago: University of Chicago Press, 2005

[Miura 2005] Miura, A. and Shinohara, K. Social intelligence design in online chat communication: a psychological study on the effects of "congestion", AI & Soc (2005) 19: 93–109.

[Mohammad 2007a] Mohammad, Y. F. O. and Nishida, T. NaturalDraw: Interactive Perception Based Drawing for Everyone, in Proc. 2007 International Conference on Intelligent User Interfaces (IUI 2007), pp. 251-260, 2007

[Mohammad 2007b] Mohammad, Y. F. O. and Nishida, T. Intention through Interaction: Toward Mutual Intention in Real World Interactions, IEA/AIE 2007, Kyoto, Japan (to be presented)

[Nijholt 2001] Nijholt, A. From Virtual Environment to Virtual Community, in Springer-Verlag LNAI 2253 New Frontiers in Artificial Intelligence (Joint JSAI 2001 Workshop Post-Proceedings), 2001.

[Nijholt 2006a] Nijholt, A. and Nishida, T. Social intelligence design for mediated communication, AI & Soc (2006) 20: 119–124, 2006.

[Nijholt 2006b] Nijholt, A., Op Den Akker, R., and Heylen, D. Meetings and Meeting Modeling in Smart Environment, Soc (2006) 20, 2006.

[Nishida 2001] Nishida, T. Social Intelligence Design -- An Overview, in: Takao Terano, Toyoaki Nishida, Akira Namatame, Yukio Ohsawa, Shusaku Tsumoto, and Takashi Washio (eds): Exploring New Frontiers on Artificial Intelligence - Selected Papers from the First International Workshops of Japanese Society of Artificial Intelligence -, Lecture Notes on Artificial Intelligence LNAI2253, Springer Verlag, December 2001.

[Nishida 2006] Nishida, T., Terada, T., Tajima, T., Hatakeyama, M., Ogasawara, Y., Sumi, Y., Xu Y., Mohammad, Y.F.O., Tarasenko, K., Ohya, T., and Hiramatsu, T., Towards Robots as an Embodied Knowledge Medium, Invited Paper, Special Section on Human Communication II, IEICE Trans. Information and Systems, Vol. E89-D, No. 6, pp. 1768-1780, June, 2006.

[Ohguro 2001] Ohguro, T., Kuwabara, K., Owada, T., and Shirai, Y. FaintPop: In Touch with the Social Relationships, in Springer-Verlag LNAI 2253 New Frontiers in Artificial Intelligence (Joint JSAI 2001 Workshop Post-Proceedings), pp. 11-18, 2001.

[Ohmoto 2006] Ohmoto, Y., Ueda, K., and Ohno, T. The real-time system of measuring gaze direction and facial features, and applying the system to discriminating lies by diverse nonverbal information, presented at SID-2006.

[Ohya 2006] Ohya, T., Hiramatsu, T., Yu, Y., Sumi, Y., and Nishida, T. Towards Robot as an Embodied Knowledge Medium—Having a robot talk to humans using nonverbal communication means, presented at SID-2006.

[Reeves 1996] Reeves, B and Nass, C. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places, CSLI/Cambridge University Pres, 1996.

[Rienks 2006] Rienks, R., Nijholt, A. and Barthelmess, P. Pro-active Meeting Assistants: Attention Please!, presented at SID-2006.

[Rosenberg 2005] Rosenberg, D., Foley, S., Lievonen, M., Kammas, S., and Crisp, M. J. Interaction spaces in computer-mediated communication, AI & Soc (2005) 19: 22–33.

[Sacks 1974] Sacks, H., Schegloff, E., & Jefferson, G. A simplest systematics for the organization of turn-taking for conversation. Language, 50 (4), 695-737, 1974.

[Saito 2005] Saito, K., Kubota, H., Sumi, Y., and Nishida, T., Analysis of Conversation Quanta for Conversational Knowledge Circulation, in: R. Khosla et al. (Eds.): *KES 2005*, LNAI 3683, pp. 296-302, 2005.

[Xu 2007] Xu, Y., Ueda, K., Komatsu, T., Okadome, T., Hattori, T., Sumi, Y., and Nishida, T. WOZ experiments for understanding mutual adaptation, AI & Society, to appear.

[Yamashita 2002] Yamashita, K, and Nishida, T. SIQ (Social Intelligence Quantity): Evaluation Package for Network Communication Tools, APCHI 2002 -- 5th Asia Pacific Conference on Computer Human Interaction - Beijing, China, 1-4 November 2002.

[Yin 2007] Yin, Z. and Fruchter, R. I-Dialogue: Information Extraction from Informal Discourse, AI & Society (to appear).