SID 04

Social Intelligence Design 2004

CTIT Proceedings of the third
Workshop on Social Intelligence Design

Anton Nijholt and Toyoaki Nishida (eds.)

CIP GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Nijholt, A., Nishida, T.

Social Intelligence Design 2004
Proceedings Third Workshop on Social Intelligence Design
A. Nijholt, T. Nishida (eds.)
Enschede, Universiteit Twente, Faculteit Informatica

ISSN 0929–0672
ISBN 90-75296-12-6
CTIT Workshop Proceedings Series WP04-02

trefwoorden: collective behavior, collaboration tools, natural interaction, multi-party interaction, community software, social agents.

Book orders:
Ms. C. Bijron
University of Twente
Dept. of Computer Science
P.O. Box 217
NL 7500 AE Enschede
tel: +31 53 4893680
fax: +31 53 4893503
Email: bijron@cs.utwente.nl
Druk- en bindwerk: Reprografie U.T. Service Centrum, Enschede

Preface

This volume of the CTIT proceedings series contains the proceedings of the third international workshop on Social Intelligence Design (SID) 2004. In this workshop we consider Social Intelligence as the ability for people to relate to, understand and interact effectively with others. Our particular concern is how SI is mediated through the use of new technologies.

The main themes of SID 2004 are:

Communities - covering community media, communication patterns in online communities, knowledge-creating, network and anonymous communities.

Collaboration Technologies and tools - covering innovations to support interactions within communities, covering a range from knowledge sharing systems, multi-agent systems and interactive systems.

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 artifacts, socially intelligent robots.

Application Domains - including design, workspaces, education, e-commerce, entertainment, digital democracy, digital cities, policy and business.

This is the third workshop on the subject of social intelligence design focused on the significance of information technology in our lives, work, home, and on the move. We are grateful to the program committee of SID 2004, the extra referees that supported their work and the invited speakers: Stefan Agamanolis (Media Lab Europe, Dublin, Ireland), Kerstin Dautenhahn (University of Hertfordshire, United Kingdom), Will Harvey (THERE Inc., Menlo Park, USA) and Stacy Marsella (CARTE Marina del Rey, USA). Obviously, we should mention also all presenters and authors of the selected papers that appear in this volume.

The first workshop, held in Matsue, Japan in 2001, opened up key themes in this area, namely, how new technologies mediate human communication and collaboration across geographical and cultural divides. Selected papers from this first SID workshop of the Japanese Society of Artificial Intelligence were published in Lecture Notes on Artificial Intelligence LNAI 2253, Springer Verlag, December 2001. The second workshop on social intelligence design was held at the university of London in Egham, United Kingdom in 2003. The papers presented there are under review for post-proceedings and a special issue of a journal.

SID 2004 has been made possible with financial support from NWO (Netherlands Organisation for Scientific Research), SIKS (Dutch research school for Information and Knowledge Systems), IOP-MMI (Senter, Ministry of Economic Affairs) and CTIT (Centre of Telematics and Information Technology). SID 2004 has been organized under the auspices of CTIT's Special Research Objective NICE (Natural Interaction in Computer-mediated Environments). We are grateful to all these supporting organizations.

At the University of Twente several people were involved in the organization of this workshop. There is already a long tradition of organizing international workshops. For example, more than twenty workshops on language technology and human computer interaction have been organized in the past ten years. And, as usual, Charlotte Bijron, Alice Vissers and Lynn Packwood helped with organizational and financial matters for this Social Intelligence Design workshop. Hendri Hondorp acted again as the technical editor for proceedings. The research institute (CTIT) gave permission to publish the proceedings in her CTIT Proceedings series. The organizers of this workshop want to thank all those who have helped in organizing this workshop.

Anton Nijholt, Toyoaki Nishida, Duska Rosenberg and Renata Fruchter July 2004

Previous SID workshops

Previous SID workshops were

SID 2001 (Matsue, Japan)	JSAI-Synsophy International Workshop on Social Intelligence Design. *May 21-22, 2001*.
SID 2003 (Egham, United Kingdom)	S.I.D. 2003: Social Intelligence Design International Conference. *July 6-8, 2003*.

List of Referees SID2004

Rieks op den Akker (University of Twente, Enschede)
Christoph Bartneck (University of Eindhoven)
Margit Biemans (Telematica Instituut, Enschede)
Betsy van Dijk (University of Twente, Enschede)
Simon Foley (Royal Holloway University, London)
Eelco Herder (University of Twente, Enschede)
Anzela Huq (Royal Holloway University, London)
Mirja Lievonen (Royal Holloway University, London)
Chris Miller (University of Southern California, Marina del Reye)
Klaas Sikkel (University of Twente, Enschede)
Hannes Vilhjálmsson (University of Southern California, Marina del Reye)

Program Committee SID2004

Anton Nijholt (Chair,University of Twente, the Netherlands)
Toyoaki Nishida (Co-Chair, Tokyo University, Japan)
Renata Fruchter (Stanford University, USA)
Duska Rosenberg (University of London, United Kingdom)
Christos Bouras (University of Patras, Greece)
Kerstin Dautenhahn (University of Hertfordshire, United Kingdom)
Tom Erickson (IBM T.J. Watson Research Center, New York, USA)
Dirk Heylen (University of Twente, the Netherlands)
W. Lewis Johnson (CARTE, Marina del Rey, USA)
Matthias Rauterberg (University of Eindhoven, the Netherlands)
Hans Schlichter (Technische Universität München, Germany)
Yasuyuki Sumi (Kyoto University, Japan)
Asako Miura (University of Osaka, Japan)
Henri ter Hofte (Telematica Institute, the Netherlands)

SID2004 Organisation Committee

Charlotte Bijron and Alice Vissers (Program Committee Secretary)
Hendri Hondorp (Proceedings Tech. Editor)

Sponsors

[1]

[2]

[3]

[4]

[1] http://www.ctit.utwente.nl
[2] http://www.nwo.nl
[3] http://www.siks.nl
[4] http://www.iop.nl

Contents

Invited Speakers

Regular Talks

Invited Talks

Social Intelligence Design 2004

Sports over a Distance, Portraits with a Sense of their Personal Space, and Telephones crossed with Flotation Tanks

Stefan Agamanolis
Media Lab Europe
The Human Connectedness Research Group
Dublin, Ireland
http://www.medialabeurope.org/hc
stefan@media.mit.edu

Abstract

Humans have a fundamental need for contact with other humans. Our interactions and relationships with other people form a network that supports us, makes our lives meaningful, and ultimately enables us to survive. The Human Connectedness research group explores the topic of human relationships and how they are mediated by technology. Our mission is to conceive a new genre of technologies and experiences that allow us to build, maintain, and enhance relationships in new ways. We also aim to enable new kinds of individual bonds and communities that were not possible before but may be beneficial or fun. In this talk, I will give a survey of the projects the group is working on in addition to an overview of Media Lab Europe, and I will outline some themes and trends to watch for in this domain of media technology.

Bio

Stefan Agamanolis is a principal research scientist at Media Lab Europe, the European research partner of the MIT Media Laboratory. He directs the Human Connectedness group, which explores the future of human relationships as mediated by technology. His research interests include remote collaboration and awareness, media spaces, automated video editing, ubiquitous computing, ambient media, sports over a distance, responsive environments, interactive storytelling, and technologies for cultural exchange.

Stefan holds MS and PhD degrees in Media Arts and Sciences from the Massachusetts Institute of Technology. Earlier in his life, Stefan studied computer science, philosophy, and film at Oberlin College. He was born and raised in the state of Ohio in the United States.

Do We Need Social Robots? A Critical Perspective

Kerstin Dautenhahn
Adaptive Systems Research Group
University of Hertfordshire, UK
http://homepages.feis.herts.ac.uk/~comqkd/
K.Dautenhahn@herts.ac.uk

Abstract

Present day robots, and even more so robots envisaged for the future, can perform different roles in human society: pets, assistants (e.g. in the area of assistive technology), guides (e.g. in museums), or servants that perform very specific functions (e.g. delivering meals). Depending on the particular function of the robot and its role in human society, social skills might or might not be desirable.

In my talk I will discuss many of such roles, and requirements regarding social skills that robots need to possess in different application areas. I will then focus on one particular role, namely the role of robots as mediators of human-human communication. Mediating human communication has been studied widely in computer based systems (e.g. CSCW, or Collaborative Virtual Environments), but when applied to robots we find that the metaphor of a "medium" is too limited: the robot as an embodied, physical and interactive entity is sharing a temporal and spatial context with its human interactants, the robot serves as an agent, a mediator, that participates in human-human communication. I will give an example of how the role of a robotic mediator has been studied in the Aurora project which investigates the use robots in therapy and education of children with autism.

This research gives a clear example where the primary role of robots is not to replace human functions, but to empower people, by mediating human contact between children with autism and other people. Note, that in this approach the emotional and intersubjective meaning of social contact is created by people in interactions, not by the robot. While in entertainment and possibly other contexts robots that express and perceive emotions might be desirable and scientifically interesting from the point of view of modelling human emotions via synthesizing artefacts, caution should be applied in how to use robotic artefacts in areas involving vulnerable subjects with core deficits exactly in the domain of perceiving and processing social and emotional information. Similar caution might be necessary in other application areas involving people with specific social and emotional needs, e.g. when designing robots as companions for elderly people, or as playmates for very young children. Generally, socially intelligent robots might show aspects of human-style social intelligence, but can hardly be faithful substitutes for human contact.

Social intelligence design for robots is a scientifically highly challenging and exciting area. However, while the field has been moving from purely laboratory prototypes to commercially available artefacts, discussions are needed on what roles exactly these robots should play in human society, and how we ensure that human values and intersubjective meaning provided by human contact will be reflected in social robotics research.

Bio

Kerstin Dautenhahn is working as a professor of Artificial Intelligence in the Department of Computer Science (Faculty of Engineering and Information Sciences) at University of Hertfordshire. In this department she is a coordinator of the Adaptive Systems Research Group. Main areas of

her research are Socially Intelligent Agents and Artificial Life. More information about it can be found on her homepage.

Kerstin is a former member of the Department of Biological Cybernetics at University of Bielefeld, Germany, 1990-1993, and AI-Lab at GMD, Sankt Augustin, Germany, 1993-1996, and VUB Artificial Intelligence Laboratory, Brussels, Belgium, until end of 1996. From January 1997 to April 2000 she is Lecturer, Department of Cybernetics at University of Reading, United Kingdom.

Selected References

Aurora project. URL: `http://www.aurora-project.com/` (Last accessed 6th April, 2004)

Dautenhahn K. (2003). Roles and Functions of Robots in Human Society - Implications from Research in Autism Therapy, Robotica, volume 21, pp. 443-452.

Dautenhahn K. (1998). The Art of Designing Socially Intelligent Agents - Science, Fiction, and the Human in the Loop, Applied Artificial Intelligence, Vol 12, 7-8, October-December, pp 573-617.

Fong T., I. Nourbakhsh and K. Dautenhahn (2003). A Survey of Socially Interactive Robots, Robotics and Autonomous Systems 42(3-4), 143-166.

Robins B., K. Dautenhahn, R. te Boekhorst, A. Billard (2004). Effects of repeated exposure to a humanoid robot on children with autism. In S. Keates, J. Clarkson, P. Langdon and P. Robinson (Eds.) *Designing a More Inclusive World*, Springer Verlag, London, pp. 225-236.

Online Socializing with Avatars

Will Harvey
THERE Inc.
165 Jefferson Drive
Menlo Park, CA 94025, USA

Abstract

Science fiction authors have been writing about virtual worlds and avatars for twenty years, and entrepreneurs have been chasing that vision for almost as long, looking for the elusive killer app. Could online socializing finally be it? Online socializing has become a mainstream phenomenon, but remains predominantly a medium of typing text. Is text the end of the line, or is online socializing destined to become a medium in which people can communicate and interact as avatars?

In the six years of developing There, we attempted to create an online world where communicating as avatars felt natural and fun, like communicating in the real world. We found the challenge to be much more difficult and subtle than we ever imagined.

This talk chronicles the challenges we had to overcome, the solutions we found, and the lessons we learned, giving examples from some of the earliest prototypes to the live service that is running today.

Bio

Will is the Founder of There, which operates a massively multi-user, graphical 3D online world for hanging out and meeting new friends. There launched commercially in October of last year and now has a vibrant community of users and economy from users who run businesses making and selling virtual clothes, conducting tours, hosting dune buggy races, throwing parties, and participating in a vast number of other events and activities.

Prior to founding There in 1998, Will ran the dynamic media products at Adobe Systems, including AfterEffects and Adobe Premiere, the world's leading video editing program. Will came to Adobe when Adobe acquired Will's previous company, Sandcastle, which Will had founded to develop network technology to enable low latency interaction over the internet. Prior to Sandcastle, Will served as Vice President of Engineering at Rocket Science Games in San Francisco, where he led the company's transition from full motion video based games to games focused on interactivity.

Prior to Rocket Science, Will founded and ran several successful game development companies while simultaneously earning his Bachelor's, Master's and Doctorate degrees in computer science from Stanford. Will's doctoral thesis introduced several important search algorithms which are now used commercially in manufacturing scheduling. Will's game companies produced Platinum and Gold game titles including Zany Golf, Immortal, and Music Construction Set, with combined sales of over a million units. Will has filed 5 patents related to networking, graphics, and automated scheduling. He wrote his first commercial video game at the age of 15. For a conversation with Wil Harvey in the journal ACM Queue - Tomorrow's Computing Today visit: `http://www.acmqueue.com/modules.php?name=Content&pa=showpage&pid=112`

Modeling Emotion in Virtual Humans: Extended Abstract

Stacy C. Marsella
University of Southern California
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90292 USA
marsella@isi.edu

Abstract

Emotions play a key role in mediating human social relationships (Davidson et al. 2003). Our interactions with each other are a source of many emotions and we have developed a range of behaviors that can communicate emotional information as well as an ability to recognize and be influenced by the emotional arousal of others. The goal of our research is to create a general and realistic computational model of emotion that accounts for emotion's cognitive sources and influences (Marsella and Gratch 2003).

Although such models can ideally inform our understanding of human behavior, we see the development of computational models of emotion as a core research focus that will facilitate advances in the large array of computational systems that model, interpret or influence human behavior. On the one hand, modeling applications must account for how people behave when experiencing intense emotion including disaster preparedness (e.g., when modeling how crowds react in a disaster Silverman 2002), training (e.g., when modeling how people respond in combatGratch and Marsella 2003), and even macro-economic models (e.g., when modeling the economic impact of traumatic events such as 9/11 or the SARS epidemic). On the other hand, many applications presume the ability to correctly interpret the beliefs, motives and intentions underlying human behavior (such as tutoring systems, dialogue systems, mixed-initiative planning systems, or systems that learn from observation) and could benefit from a model of how emotion motivates action, distorts perception and inference, and communicates information about mental state. Emotions play a powerful role in social influence, a better understanding of which would benefit applications that attempt to shape human behavior, such as health interventions (Marsella et al. 2003), psychotherapy applications (Rothbaum et al. 1999), tutoring systems (Lester et al. 1999; Shaw et al. 1999), and marketing applications (André et al. 2000; Cassel et al. 2000). Lastly, models of emotion may give insight into building models of intelligent behavior in general (e.g., Damasio 1994; Minsky 1986; Oatley and Johnso-Laird 1987; Simon 1967; Sloman and Croucher 1981).

In particular, computational models of emotions have become important in the development of social training environments, an environment where people can explore stressful social interactions in the safety of a virtual world. Social training environments are often populated with virtual humans, realistic human-like software artifacts that not only look and act like people but also can interact with them. Given the focus on stressful social interactions, emotion modeling is a central component in the design of virtual humans for these environments.

Our research into emotion for virtual humans has lead to EMA (**Em**otion and **A**daptation), a computational model of emotion and coping. In this talk, I will discuss EMA, the underlying motivations that lead to its development and the directions the research is going.

Keywords: Models of Emotion, Virtual Humans

References

André, E., Rist, T., Mulken, S. v., & Klesen, M. (2000). The Automated Design of Believable Dialogues for Animated Presentation Teams. In J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), Embodied Conversational Agents (pp. 220-255,). Cambridge, MA: MIT Press.

Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., & Yan, H. (2000). Human conversation as a system framework: Designing embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), Embodied Conversational Agents (pp. 29-63). Boston: MIT Press.

Damasio, A. R. (1994). Descartes' Error: Emotion, Reason, and the Human Brain. New York: Avon Books.

Davidson, R. J., Scherer, K., & Goldsmith, H. H. (Eds.). (2003). Handbook of Affective Sciences. New York: Oxford University Press.

Gratch, J., & Marsella, S. (2003). Fight the way you train: the role and limits of emotions in training for combat. Brown Journal of World Affairs, X(1)(Summer/Fall).

Lester, J. C., Stone, B. A., & Stelling, G. D. (1999). Lifelike Pedagogical Agents for Mixed-Initiative Problem Solving in Constructivist Learning Environments. User Modeling and User-Adapted Instruction, 9(1-2), 1-44.

Marsella, S. & Gratch, J. (2003). Modeling coping behaviors in virtual humans: Don't worry, be happy. Paper presented at the Second International Joint Conference on Autonomous Agents and Multi-agent Systems, Melbourne, Australia.

Marsella, S., Johnson, W. L., & LaBore, C. (2003). Interactive Pedagogical Drama for Health Interventions. AIED 2003, 11th International Conference on Artificial Intelligence in Education, Australia.

Minsky, M. (1986). The Society of Mind. New York: Simon and Schuster.

Newell, A. (1990). Unified Theories of Cognition. Cambridge, MA: Harvard University Press.

Oatley, K. & Johnson-Laird, P. N. (1987). Cognitive Theory of Emotions. Cognition and Emotion, 1(1).

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., & Swartout, W. (2002). Toward a New Generation of Virtual Humans for Interactive Experiences. IEEE Intelligent Systems, July/August, 32-38.

Rothbaum, B. O., Hodges, L. F., Alarcon, R., Ready, D., Shahar, F., Graap, K., et al. (1999). Virtual Environment Exposure Therapy for PTSD Vietnam Veterans: A Case Study. Journal of Traumatic Stress, 263-272.

Shaw, E., Johnson, W. L., & Ganeshan, R. (1999). Pedagogical Agents on the Web. Paper presented at the Proceedings of the Third International Conference on Autonomous Agents, Seattle, WA.

Silverman, B. G. (2002). Human Behavior Models for Game-Theoretic Agents: Case of Crowd Tipping. CogSci Quarterly, Fall.

Simon, H. A. (1967). Motivational and emotional controls of cognition. Psychological Review, 74, 29-39.

Sloman, A., & Croucher, M. (1981). Why robots will have emotions. Paper presented at the International Joint Conference on Artificial Intelligence, Vancouver, Canada.

Regular Talks

Social Intelligence Design 2004

Supporting Internalization Process with Wet Community Awareness, Ubiquitous and Weblog

Shintaro Azechi
School of Business Administration, Asahi University
Hozumi 1851, Mizuho, Gifu, JAPAN 501-0296
az@pvq.jp

Abstract

The community based aspect of regards Social Intelligence Design as total community activities for knowledge creation. In such activities, knowledge circulative changing, tacit mode and formal mode are important. Although many knowledge creating methods and community mediating system are available for externalization process that changes knowledge from tacit to formal, there rarely exists the method or system for internalization process that changes knowledge from formal to tacit. This article introduces applications of Informational Humidity Model for supporting knowledge creating community. Especially, it is introduced how the wet community support the internalization process, using awareness system, ubiquitous system and weblog system.

Keywords: Knowledge Creating Community, Informational Humidity Model, Wet Community, Awareness, Ubiquitous, Weblog

1 Introduction

What and how the method and system can support the knowledge creating process? It is one of important aspect of social Intelligence Design (SID).

Sharing tacit knowledge is important for knowledge creation. However, especially in Japan where technological generation gap has been increasing mainly because of decreasing of young population, it is not well done in many cases.

In this paper, the problems and solving methods and systems how the tacit knowledge should be shared in a community. At first, the knowledge creation process in community is briefly described. The second, the *Informational Humidity Model* is introduced. The third, three systems and services on line are introduced which are thought to useful for mediating *wet information* and tacit knowledge. At last, needed features for to realize the methods for sharing tacit knowledge is discussed.

2 Knowledge Creation Process

Although there are many famous knowledge creating techniques, most of them focus on externalizing process, how to change individual tacit knowledge to shared formal process. Similarly, community mediating systems also support mainly externalization process of collaborative knowledge creation. For example, anonymous bulletin board systems or Public Opinion Channel (Nishida, 1999), which are community supporting systems, make members easily to expressing their tacit knowledge to explicit publicly.

On the other hand, internal process is not supported sufficiently. A rare case is Akao's collaborative idea finding technique 'Reviewed Dendrogram' (Akao, 1990) that makes and arranges accumulated question and answer cards for internalization process. Unfortunately, however, there is no supporting system for this method.

Nonaka (Nonaka, 1995) explained knowledge creating process as a circulation of mode of knowledge in a community. In externalization process, community member's individual tacit knowledge

comes to express to other community members and changed to shared formal knowledge. On the other hand, in internalization process, created knowledge by interaction comes into individual memories and holed for the next chance to express as formal knowledge.

If there is no idea for supporting method or system for internalization process, the ring of knowledge creating process would be missed. How can solve this serious problem?

Recently, the role of *Community of Practice* (Wenger et al., 2002) is emphasized. In a community of practice, people learn other members tacit knowledge through their experience. Wenger emphasizes that for innovation and enhancing continual growing of industry, it is important to mediate project members by well making of community of practice. For example, co-experiences of continuum project or sharing a part of life would be useful for this purpose. Co-working of older and younger persons would also makes sharing of tacit knowledge between generations, and it will contribute to future growth of the organization.

Wenger describes the merits of well made communities of practice. However there are few ideas for how to support or make community of practice with networking system and how enhance sharing of tacit knowledge. One important perspective is that community of practice would be mediated with human relationship, emotional ties and trust. From this view point, mediating personal information for example benefit to self disclosures are needed to well making of community practice. In the next section, a model is introduced to lead how the personal information would be mediated in a community, the Informational Humidity Model.

3 Informational Humidity Model

Azechi (Azechi, in press) described Informational Humidity Model (IHM), a classification of information in community, wet versus dry information. the characteristics of the community, wet versus dry community. The dry information is the core part of the message, which contains the topic being discussed. On the other hand, wet information contains attributes of the message sender. This dual structure of a message and its role in a persuasive situation is supported by cognitive psychological evidence. It is thought as one of fundamental classification of message exchanged in communication.

Azechi also identified community, dry versus wet. In dry community, members cannot specify or identify other member's personal information. On the other hand, wet community where members specify other member's personal information sufficiently. Azechi concerned that the dry community is suitable for externalization process. There are no barriers to prevent express information, for example authoritative power from another member in a dry community. On the other hand, wet community is suitable for internalization process. Plentiful personal (wet) information is assumed supporting to know other member's what to know in the wet community.

Preparing plentiful wet information in a community is considered to supporting well making of community of practice. If a person can get other community members easily and immediately, the person would be able to make good relationship with the member soon. It is for example like as much making in romantic relationship. If a couple want to meet and make emotional ties, knowing personal information mutually should be important premise. Supporting plenty supplying or mediating wet information is useful for persons know well each other. In community of practice, both romantic and business, supporting wet information mediating would enhance doing well the community.

In a wet community where wet information is plenty mediated, that would be a well constructed community of knowledge, tacit knowledge of members would be well shared. And also, internalization process of knowledge creating would be enhanced. Based on IHM, the next section describes systems that support wet community for internalization process.

4 Wet Community with Mediating Systems

Wet community is defined as a place in which members can know other member's personal information plentifully. In general, real community has the characteristics of the wet community; however mediated community is rather dry because media tend to fail to transfer wet information.

However some trials of online systems and services are existed. Especially, three tools seem to be suitable for mediating wet information and constructing wet community; system to supply awareness, ubiquitous systems and weblog communities. It is explained how those systems could make wet community and enhance external process of knowledge creation in collaborative way.

4.1 Awareness System

Supplying awareness has been a key subject of the study for Computer Supported Collaborative Works (CSCW). In this research area, awareness supply makes communication of people who are mediated with the system, and it enhances the quality of intelligent products of the CSCW. It is supposed that awareness is a constructive part of wet information and it make collaborative knowledge creation process (especially internal phase) smoothly.

For one example, Azechi et al. (Azechi, 2003) implemented awareness supplying system and elaborated it in the educational situation setting. Their system is a kind of live camera broadcasting teacher's behavior in her/his office on the web page, together with her or his day / week schedule and status on the time [1]. Azechi et al. found limited effect of the system for students who are let to write a brief report for a psychological topic. A tendency was observed that system support students understanding of topic in both remote and neighboring situations, otherwise without the system, remote student's understanding are insufficient.

How and why awareness supplying system can support internalization process should be discussed with empirical data in future work. And also IMH should describe the cognitive and group mechanism how and why awareness supplying enhances internalization process.

4.2 Ubiquitous

Recently ubiquitous environment has been developed owing to the popularization of mobile phones. Ubiquitous environment is defined as the situation in which anyone can connect to the net and communicate with other persons anywhere anytime. Actually, perfect ubiquitous environment have not been realized, however pseudo-ubiquitous or future seeds are appeared. Although the service of browser-phone and hot spot are expensive or area limited, how use ubiquitous environment comes to be centered in business or educational research area.

With ubiquitous environment, wet information could come to be supplied fluently because one can send her or his idea or images to the net immediately from any where. Unless the system is implemented as anonymous media, ubiquitous environment empower the function of wet information, strengthening human relationships and group maintenance.

And also ubiquitous environment is suitable for educational purpose with wet information. Students can express their findings about their subject to the net mediated system anytime. This easy sharing of formal knowledge with other students is an effective way of knowledge internalization in a educational community like a classroom. Though it is also said to describe how and why the wet information and ubiquitous system leads the process is needed, this aspect would be possible to solve how the system help internalization process.

4.3 Weblogs

A weblog is generally defined as a website in which people write everyday events and thoughts, or make links and comments to other websites that is renewed almost everyday. Because many weblog tools support to make it easy in recent, the size and members of weblog communities come to be increased. Most of weblog contains wet information, not directly appears writers personal information, but shows personality indirectly. Most weblog writers use handle and do not shows her or his real sociality (name, age, race, occupation and so on), but their summed contents of articles shows what person s/he is; her or his personality.

Weblog also seems effective to internalization process of knowledge creation. When a writer makes an article that contains some of knowledge, it would come to accumulate on the web as collective tacit knowledge and wait to be search as new knowledge (externalization). This

[1] A part of this system can be browsed on http://pvq.jp/notice/

process of weblog seems to be concentrated to self-presentation and motivation for informational expression. Unless there is no motivation to express self, renewals of articles may come to be over soon. This motivational process seems to be concerned with the function of wet information; however it should be needed to examine more detailed.

5 Perspectives

This paper shows a possibility to solve a problem for to enhance knowledge creation in community, which is knowledge internalization. Making community of practice is an answer. For answering the question how the community of practice is made or supported by networking system, the IHM is proposed.

What features are needed for to achieve this purpose? There are three possibilities. One is extending the bandwidth of the net and mediating video and audio information. Although this would seem very good and easy perspective, people should not use this willingly. It misses main benefit of networking tools, visual anonymity. With video and audio system, people cannot use creative self, present her of his own ideal self image efficiently.

The second, making easy to express slight information to the community seems useful. If people externalize and internalize knowledge anytime, sharing tacit knowledge should be enhanced. However it is remarkable for human cognitive ability to express information. Because it would be pressure for human cognitive resources, some supporting feature is needed.

At last, accumulation and publication of personal documents is available. This would be a good wet material to know how the person has personality. However this should be made valance to privacy protections.

The three examples of present system, awareness providing, ubiquitous, weblog present the future possibility of the features of wet community supporting system. In the near future, how the networking tool support sharing tacit knowledge and, for the purpose, mediating personal information would be important focus of knowledge creating supporting system. For the development of system for this purpose, the analysis in this paper should be available.

References

Akao, Y. (1990). Quality Function Deployment: Integrating Customer Requirements into Product Design, Productivity Press Inc.

Azechi, S. (in press). Informational Humidity Model: Explanation Dual Modes of Community for Social Intelligence Design, *AI and Society*.

Azechi, S., Miura, A., Shinohara, K. and Machida, K. (2003). The Effects of Digital Awareness for e-Learning (Japanese), *Transactions of Japanese Society for Information and Systems in Education*, 20, 2, 227-233.

Nishida, T., Fujihara, N., Azechi, S., Sumi, K., and Hirata, T (1999). Public Opinion Channel for Communities in the Information Age, New Generation Computing, 17, 417-427.

Nonaka, I., and Takeuchi, H. (1995). The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press.

Wenger, E., McDermott, R. and Snyder, W. M. (2002): Culrivating Communities of Practice, Harverd Business School Press.

Measuring Affective Benefits and Costs of Awareness Systems Supporting Intimate Social Networks *

Joy van Baren, Wijnand IJsselsteijn, Panos Markopoulos, Natalia Romero
Eindhoven University of Technology
P.O.Box 513
5600 MB Eindhoven
The Netherlands
{J.K.v.Baren|W.A.IJsselsteijn|P.Markopoulos|N.Romero}@tue.nl

Boris de Ruyter
Philips Research Eindhoven
Prof. Holstlaan 4
5656 AA Eindhoven
The Netherlands
boris.de.ruyter@philips.com

Abstract

Human beings are social beings. We have a fundamental need to communicate - to form, maintain and enhance social relationships. Members of intimate social networks (e.g. family members, close friends) that live apart, stay in touch with each other through a range of synchronous and asynchronous communication media. This paper describes research focused at developing and testing a questionnaire measure that is sensitive to the affective benefits and costs that communication media may have for their users. First, a requirements study is reported, investigating real-life communication between family members, in order to identify their various communication needs and patterns, as well as their use of various existing communication media. We subsequently describe the development of a new questionnaire measure, the Affective Benefits and Costs in Communication Questionnaire (ABC-Q), that is aimed to assess the affective characteristics of communications, which were hitherto largely ignored in measures of social presence or communication effectiveness. We conclude this paper by describing a field study, evaluating the ASTRA awareness system that is specifically aimed at supporting intimate social networks. During this study the ABC-Q was first applied. Results in terms of reliability and validity of the questionnaire show that the ABC-Q is a promising measure.

Keywords: Affective Benefits, Awareness Systems, Communications Systems, Questionnaire Development, User-Centred Design

1 Introduction

The interactions and relationships we have with other people form an essential social network that supports us and adds meaning to our lives. This well-known fact is illustrated by the massive success of recent communication media such as email, mobile telephony, and SMS, but the basic insight can be traced back to the days of Aristotle, or even earlier. Maslov's theory of human needs, formulated in the 1950s, illustrates that social interaction is essential to satisfying human needs at several levels, in particular needs for belonging, love, and esteem, although even at the

*The work reported in this paper has been supported by the ASTRA project (IST-2001-39270) funded through the IST programme of the EU, FET Open scheme.

more basic levels of physiological and safety needs, communication and co-ordination is essential to mental and physical wellbeing, and ultimately, survival.

Reflecting this sentiment, William Schutz (Schutz, 1966 in Washington (2001)) describes two basic human needs in his interpersonal needs theory:

- Inclusion: the need for the company of others to feel accepted.
- Affection: the need to form emotionally close relationships.

In other words, we need the company of others in order to feel worthwhile and understood, and to express feelings such as friendship and love. The forming, maintaining and enhancing of relationships with other people is one of the most powerful drives for people to engage in communication. Traditionally, these needs have been addressed in the most natural way: face-to-face. Due to changing lifestyles and the emergence of new technologies, nowadays many of our daily interactions have become technologically mediated. Convenient as it may be when distance or time limit the opportunities for real meetings, there are severe doubts whether mediated communication can afford the same affective characteristics as face-to-face communication.

Most existing communication media have been developed to support the exchange of information rather than emotion. Often people engage in communication not merely to exchange information, but to pursue social and emotional goals such as being reassured about the well-being of others or to let other people know you are thinking about them. In this case, the informational content of a message is of secondary importance to the emotional, relational content that is being transmitted. Kuwabara et al. (2002) refer to this distinction as contents-oriented communication versus connectedness-oriented communication. Liechti and Ichikawa (1999) introduced the concept of affective awareness, defined as "a general sense of being close to one's family and friends". We would like to introduce a related concept, *connectedness*, which we define as "a positive emotional appraisal which is characterized by a feeling of staying in touch within ongoing social relationships" (IJsselsteijn et al., 2003).

A class of systems, which tries to take into account this need for connectedness, are awareness systems. Awareness systems help people to effortlessly maintain awareness of each other's whereabouts and and activities, thus facilitating lightweight, emotional, informal forms of communication. In line with Weiser and Brown (1995)'s notion of calm computing, such systems can be always-on, yet very gentle in terms of attentional demands. Pioneering examples of awareness systems include Media Spaces (Bly et al., 1993) and Portholes (Dourish and Bly, 1992), which were originally developed to serve the need to stay aware of each others presence and availability within the workplace.

Traditional assessments of communication media have often focused on social presence measurement. Social presence can be defined as the feeling of "being together" that people may experience when communicating through a medium (Short et al., 1976). Although highly relevant, such measurements do not address less direct, longer-lasting feelings resulting from communication. For example, do people think about each other more after sending or receiving an e-mail? Do they feel more involved in each other's lives if they use instant messaging on a regular basis? Do they feel connected with each other while talking through the telephone, and how long does this feeling linger on after the conversation has been ended? In order to answer these and related questions, new methods are needed. We need to identify and operationalize affective characteristics of communication, such as connectedness, which are essential for evaluating new media.

This paper describes the research surrounding the development of a new measurement instrument that is sensitive to measuring affective benefits and costs of communication. In the next section we first discuss a requirements gathering study in which we investigated real-life communication between family members, to identify their various communication needs and patterns, as well as their use of various existing communication media to support those needs. Subsequently, we describe the development of a new questionnaire measure assessing affective costs and benefits of communication. Finally, we will relate the results of a field study evaluating an awareness system aimed at supporting intimate social networks, such as families. During this study the newly developed questionnaire was first applied.

2 The ASTRA Requirements Study

The European ASTRA project (Awareness Services and Systems - towards Theory and Realisation - see: `http://www.presence-research.org/Astra/`) was specifically aimed to investigate awareness systems and services for social purposes (e.g., to help geographically separated but closely related people to stay in touch). In the context of the ASTRA project, a requirements study was conducted to explore communication needs in the home. Three family clusters, consisting in total of 13 people, took part in the complete study; 4 additional participants took part in the interviews only. By family clusters we mean individuals related with family ties, who live in more than one household. Examples of family ties in our study were parent-child, grandparent-grandchild and sister-sister. Different age groups were represented in the sample and participants had varying levels of affinity with technology.

Participants were interviewed about their attitude towards communication with family members living in another household, their experiences with existing communication means and finally their wishes for future systems. The family clusters kept a diary for one week, in which they described the time, location, device, reason, and feelings associated with each contact.

The results of the study confirm that communication with family members is highly valued. In particular, contacts for social and emotional reasons are appreciated. Participants reported that they can feel very connected with each other as a result of mediated communication. They describe this feeling as thinking about each other, feeling involved, and being emotionally touched. This feeling does not automatically disappear after the contact is ended, but may linger on for minutes, hours or even days afterwards.

Participants acknowledge and enjoy it if they perceive that another person has made an effort to contact them. The more personal this effort is, the more it is valued. For example, the effort taken to start up a PC, log in and type an e-mail is not valued as much as choosing a postcard to fit the personal taste of the receiver, and personalising it further by a handwritten message.

Apart from these positive experiences, there also appear to be costs associated with communication. Communication for social and emotional purposes can give rise to powerful expectations. Participants describe how they can feel disappointed, frustrated or even angry when these are not fulfilled. Because of this, strong feelings of obligation may arise. People are afraid to hurt or offend the other by not responding to a contact or failing to initiate one at the expected time or frequency. Timing seems to be crucial; participants indicated that they would like to share events of their everyday life with their relatives, right at the moment when they happen. On the other hand, this timing should be negotiated with the receivers so as not to disrupt them in the course of their daily life. A final cost of communication is the time and energy that it can take.

3 Development of the ABC-Questionnaire

An extensive literature review revealed that at present, there appears to exist no suitable measurement tool to address the affective characteristics of communication means. This prompted us to develop such a tool ourselves. A main difficulty was a lack of sufficient relevant literature to supply theoretical foundation. To generate alternative starting points, we conducted a brainstorm session with five people working in the area of awareness and communication research. Input to this brainstorm was existing literature, the ASTRA requirements study discussed earlier, and the results of an e-mail questionnaire about connectedness (Van Lanen, 2003). Each participant generated a number of aspects that he or she thought the questionnaire should address. These aspects were subsequently grouped in an affinity diagram to identify main concepts. The resulting ten main concepts were: Privacy, Obligations, Expectations, Effort, Thinking about Each Other, Situational Awareness, Connectedness, Sharing Experiences, Recognition, and Group Attraction. We identified two main dimensions: the first four concepts can be seen as Costs of communication, whereas the latter six relate to Benefits. Therefore, we named the questionnaire the *Affective Benefits and Costs in Communication Questionnaire* (ABC-Q).

For each concept except Group Attraction, a scale consisting of 6 questions was generated based on the brainstorm results. For group attraction, we found an existing questionnaire (Evans and Jarvis, 1986), from which we took the six items that were most appropriate to our situation,

and incorporated them in our questionnaire. All items have a seven-point scale, which runs from 'strongly agree' to 'strongly disagree'.

The first version of the ABC-Q was reviewed by two experts on questionnaire design. They were asked to check whether items were understandable and measured the intended concept. On their advice, two items were removed from the questionnaire because they were unsuitable, and the wording of five items was slightly changed to make them clearer. A pilot test was conducted to gather data for the item selection. 20 participants (students and employees of Eindhoven University of Technology) completed the questionnaire, which now contained 58 items. The goal was to reduce the number of items to 40, i.e. four per scale. Criteria to remove items were:

- The item did not contribute sufficiently to the scale (low/negative item-total correlation)
- Cronbach's alpha would improve when the item was removed.
- An item did not discriminate (low variance and/or extreme mean)

Reliability of the Effort scale turned out to be very low. A closer look at the inter-item correlations suggested that this scale was multidimensional. A short interview with several respondents confirmed that effort is a complex concept; it can be positive in one situation and negative in another. This observation is confirmed by the results of the requirements study, in which participants reported different types of effort. Because there was no time in the current project to adapt the effort scale, it was removed from the questionnaire and addressed instead using interviews.

After removing the Effort Scale internal consistency, as measured by Cronbach's alpha, of the ABC-Q as a whole was .86, which indicates that the questionnaire is very reliable. Values for Cronbach's alpha for the scales of the ABC-Q can be found in Table 1.

Table 1: Internal Consistency of the ABC-Q

ABC-Q	*.86*
Costs	*.64*
- *Obligations*	*.41*
- *Expectations*	*.73*
- *Privacy*	*.76*
Benefits	*.88*
- *Thinking about Each Other*	*.70*
- *Situational Awareness*	*.84*
- *Connectedness*	*.58*
- *Sharing Experiences*	*.81*
- *Recognition*	*.72*
- *Group Attraction*	*.84*

4 The ASTRA Field Test

The ABC-Q was first applied in a field test conducted in the context of the ASTRA project. In this project, a prototypical awareness system was developed to help distributed family members to stay in touch with each other by sharing moments of their daily life in the form of pictures, drawings and short text messages (Markopoulos et al., 2004). The field test aimed to evaluate how people experience the usage of this system in their daily life. This study offered an excellent opportunity to investigate whether the newly developed ABC-Q would yield sensible and stable results.

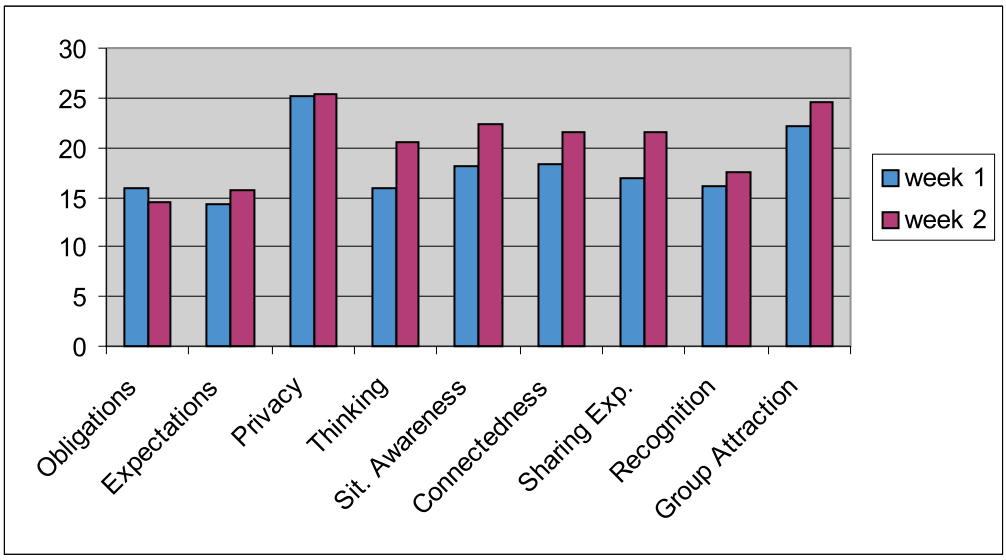


Figure 1: Mean scores on the ABC-Q in the first and the second week of the field test.

4.1 Participants

Two families, both distributed over two separate households, took part in the field test. In total, thirteen people participated, consisting of seven adults, four teenagers, and two children. The children were too young to complete the questionnaires, therefore questionnaire data were gathered from eleven participants.

4.2 Method

A within-subjects design was used, consisting of two phases that lasted a week each. In the first week, communication between two related households using existing communication means was studied. In the second week, the ASTRA system was introduced in both households. In both the first and the second week, group interviews were conducted. Participants also kept a diary and filled out two questionnaires: the IPO-SPQ measuring social presence (de Greef and IJsselsteijn, 2001), and the ABC-Q.

4.3 Results

In this paper, only results which are relevant for the evaluation of the ABC-Q will be discussed. For a more extensive description of results, please see the ASTRA Project Assessment deliverable (available from: `http://www.presence-research.org/Astra/`).

The mean scores for all scales of the ABC-Q are shown in Figure 1 (N.B. For all scales, the minimum score is 4 and the maximum score is 28). The Wilcoxon test was used to test for significant differences between scores. The results show that the first three scales, relating to Costs, do not differ significantly between the first and the second week. The Benefits scales, however, show some marked differences. Participants thought about each other more often in the second week, when they were using ASTRA ($Z = -2.67$, $p = .008$). Also, their awareness of the situation of their family members was much higher ($Z = -2.31$, $p = .021$). Participants indicated they felt more connected to each other in the second week ($Z = -2.02$, $p = .043$). They also felt they were sharing more experiences with each other ($Z = -2.38$, $p = .011$). The level of group attraction was higher in the second week ($Z = -2.23$, $p = .026$). The recognition, finally, was slightly higher in the second week but this difference was not significant.

In summary, the results from the ABC-Q imply that the ASTRA system did not increase

the costs associated with communication, but did deliver substantial affective benefits to the participants. An interview which was conducted with each of the 4 participating households after they used the ASTRA system confirmed these results. Participants reported that they felt more in touch with each other in the second week. They thought about each other more often, were more aware of what their family members were doing, and could share more of their own experiences. Many remarks were made that support this. Some illustrative quotes:

> *"There was more involvement, more curiosity. I was thinking about them much more than usual."*

> *"You become more conscious of what your family members are doing, and you also become curious."*

> *"It is so good to see what they are doing. I always want to stay in touch, but I normally don't have the time. Now it is easier."*

> *"It was fantastic to keep them up-to-date."*

5 Conclusions

The results from this study imply that the ABC-Q is a promising measure. The initial pilot testing during the development phase has shown that it is a reliable measure. In the field test, the ABC-Q proved to be sensitive to changes in condition (in our case, the introduction of the ASTRA system). Although validity has not yet been formally established, the correspondence between the ABC-Q scores and the interviews is striking and encouraging. Therefore, we believe that the ABC-Q will be a useful addition to the currently limited spectrum of relevant measurement tools in the field of communication.

6 Future work

We are currently in the process of redesigning and evaluating a new version of the ABC-Q. This version is different from the initial ABC-Q in three ways.

First of all, the concept of Effort has been brought back into the questionnaire. Originally, this was a scale in the Costs dimension, but in the pilot test it turned out to be multidimensional and therefore unreliable. As explained previously, we think the reason for this is that there are different types of effort. The first is associated with time, energy or financial costs, for example buying a stamp and walking to the post office to mail a letter. We call this "process effort". The second type of effort is meaningful to the communication, for example choosing a postcard with a nice picture on it and writing a personal message on it. We call this "personal effort". Process effort, even though time or money is invested by the sender, is not proportionally appreciated by the receiver. It can therefore be considered a Cost. Personal effort, on the other hand, is enjoyed by the sender, who has the feeling of sharing something special, and very much appreciated by the receiver. It can therefore be seen as a Benefit. Accordingly, we have added two seperate effort scales to the ABC-Q: one for process effort, which is part of the Costs dimension, and one for personal effort, which is part of the Benefits dimension.

A second change is related to the Sharing Experiences scale and the Situation Awareness scale. Questions in these two scales are quite similar, and high correlations between the two scales were found in both the pilot test and the ASTRA field test. We therefore decided to combine them into one scale.

Finally, the ABC-Q was extended with questions addressing reciprocity. In the new version, each scale addresses not only the participant's own feelings, but also the participant's perception of the feelings of his or her communication partner. For instance, the item "I feel involved in the other person's life" is now matched by the item "(I think that) the other person feels involved in my life".

We will conduct a study in the near future to evaluate the new ABC-Q. We aim to compare different media such as e-mail, traditional phone, mobile phone and SMS.

References

Bly, S., Harrison, S., and Irwin, S. (1993). Mediaspaces: Bringing people together in video, audio, and computing environments. *Communications of the ACM*, 36:29–47.

de Greef, P. and IJsselsteijn, W. (2001). Social presence in a home tele-application. *CyberPsychology and Behavior*, 4:307–315.

Dourish, P. and Bly, S. (1992). Portholes: Supporting awareness in a distributed work group. *Proceedings of ACM CHI '92*, pages 541–547.

Evans, N. and Jarvis, P. (1986). The group attitude scale. *Small Group Behavior*, 17:203–216.

IJsselsteijn, W., van Baren, J., and van Lanen, F. (2003). Staying in touch: Social presence and connectedness through synchronous and asynchronous communication media. In Stephanidis, C. and Jacko, J., editors, *Human-Computer Interaction: Theory and Practice (Part II), Volume 2 of the Proceedings of HCI International 2003*, pages 924–928. Lawrence Erlbaum, Hillsdale, NJ.

Kuwabara, K., Watanabe, T., Ohguro, T., Itoh, Y., and Maeda, Y. (2002). Connectedness oriented communication: Fostering a sense of connectedness to augment social relationships. *IPSJ Journal*, 43:3270 – 3279.

Liechti, O. and Ichikawa, T. (1999). A digital photography framework supporting social interaction and affective awareness in home communication. *Proceedings of the International Workshop on Handheld and Ubiquitous Computing HUC '99*, pages 186–192.

Markopoulos, P., Romero, N., van Baren, J., IJsselsteijn, W., de Ruyter, B., and Farschian, B. (2004). Keeping in touch with the family: Home and away with the ASTRA awareness system. *Proceedings CHI 2004, CHI Letters*, 6.

Short, J., Williams, E., and Christie, B. (1976). *The Social Psychology of Telecommunications*. John Wiley, London.

Van Lanen, F. (2003). Staying in touch over distance: An exploration of the concept of connectedness. Unpublished Masters Thesis. Eindhoven University of Technology.

Washington, W. (2001). Exploring ambient media presence awareness. Unpublished Masters Thesis. University of Washington.

Weiser, M. and Brown, J. (1995). Designing calm technology. PowerGrid Journal, 1(1). Available online: `http://powergrid.electriciti.com/1`.

Cross-Cultural Study of Expressive Avatars*

Christoph Bartneck ♣
christoph@bartneck.de

Toru Takahashi ♢ ♣
toru@atr.jp

Yasuhiro Katagiri ♢
katagiri@atr.jp

♣ Eindhoven University of Technology,
Department of Industrial Design,
Den Dolech 2,
5600 MB Eindhoven, The Netherlands

♢ ATR Media Information Science Research Labs,
2-2-2 Hikaridai, Keihanna Science City, Kyoto,
619-0288, Japan

Abstract

Avatars play an important role in international online communities. While certain simple expressions, such as facial emotional expressions are cultural independent, more complex expressions might not be. Therefore we conducted a cross-cultural study to investigate the influence of the users' cultural background (Japanese or Dutch) on their perception of avatar's expressions in terms of perceived arousal and valence. A significant gender difference was found for valence. Women and in particular Japanese women rated the expressions more positive.

Keywords: cross-culture, avatars, expressions, animations, arousal, valence

1 Introduction

Internet forums and their resulting online communities are currently becoming increasingly important (Rheingold (2002); OSDN (2004)) and help people to share problems and to organize activities. At the same time the communication in today's forums is restricted to text messages, which offers only a fraction of the information available in face-to-face conversation. This results in frequent misinterpretations especially if the messages involve irony. The widespread usage of emoticons :-) demonstrates that pure textual information lacks natural communication channels. To overcome these restrictions virtual representations of the participants, so called avatars, are being developed (Damer (1997)) The embodiment of the participants through avatars enables the participants to use body language and facial expressions in their messages. Previous studies show that avatars can play an important role in cross-cultural communication (Isbister et al. (2000); O'Neill-Brown (1997)).

The facial expression of emotions has been shown to be cultural independent (Ekman and Friesen (1971)) but some avatars , such as in TelMeA (Takahashi et al. (2003)) or Flirtboard (Krenn et al. (2004)) use more complex expressions which might be perceived differently by users with different cultural backgrounds. These complex expressions can consist of a series of movements of all components of the avatar. It might, for example, simultaneously smile and jump up and down before it waves it hands to the user. Since the internet is available in many countries it is very likely that users with diverse backgrounds will communicate which each other in these forums.

*We would like to thank Jun Hu for his helpful comments on this paper. This research was partly supported by the National Institute of Information and Communications Technology (NICT) of Japan.

We therefore conducted a study that investigated the influence of the user's cultural background on how he/she perceives the avatar's complex expressions.

The repertoires of avatars are not limited to emotional expression, but also include conversational and social cues, such nodding and spatial distance between avatars. Therefore a general two dimensional model of valence and arousal (Schlossberg (1952); Hendrix et al. (2000)) appears to be best suited to model the user's perception of the expressions.

Much work has been done on cultural research and particularly Hofstede's (Hofstede (1984)) four constructs (power distance, uncertainty avoidance, individualism, and masculinity) has been the foundation of much comparisons between cultures. A precise definition of culture and differences between cultures is not in the focus of this study. However, we are certain that the Japanese and Dutch cultures differ significantly enough to function as a factor in our experiment.

2 Methodology

To have a sufficient sample of animations a professional Japanese designer created thirty expressions of two different avatars. The animations of the two avatars were almost identical. Figure 1 shows the waving animation for both avatars.

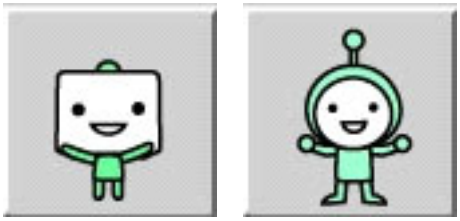


Figure 1: Example stimuli of the two avatars.

Figure 2 shows the most intense expressions of the square avatar's animations. The animations are numbered according to the order in the experiment.

2.1 Design

A 30 (animation) x 2 (avatar) x 2 (culture) within/between subjects experiment was conducted in which animations (see Figure 2) and avatars (see Figure 1) were the within factors and culture the between factor. The exact same experiment was once conducted at ATR, Kyoto, Japan and once at the Eindhoven University of Technology, The Netherlands.

2.2 Participants

In the Japanese experiment 13 men and 14 women ranging from 18 to 25 years of age participated. Participants of the Dutch experiment consisted of 16 men and 11 women ranging from 18 to 51 years of age. All participants were university students or university employees. The participants received a financial recompense for their participation.

2.3 Measurements

The participants were asked to rate the arousal and valence of the animation on six 5-point scales. The scales were anchored with a pair of antonymous words in Japanese or Dutch language. The approximate translation into English is as follows:

Arousal

- rough behavior – gentle behavior
- excited behavior – calm behavior

Figure 2: The 30 animations of the square character

- aware behavior – numb behavior

Valence

- cooperative behavior– counter-cooperative behavior
- behavior of agreement– behavior of disagreement
- permitting behavior– denying behavior

The score of arousal and valence for each animation was calculated by calculating the mean of the three answers to the arousal or valence questions.

2.4 Process

The animations and the questionnaire were presented on a computer screen in front of the participants. After a training session the participants had the opportunity to ask questions about the process of the experiment before they would begin with the experiment. The 60 stimuli were presented in random order and the participants could replay the animation as often as they wanted by pressing the "replay" button (see Figure 3). They had to mark each scale before they could press the "okay" button to continue to the next stimuli. The whole process took one hour.

3 Results

A 30 (animation) x 2 (avatar) x 2 (culture) x 2 (gender) repeated measures ANOVA was conducted. The animations had a significant effect on arousal ($F(29, 1450) = 79.51, p < .001$) and valence ($F(29, 1450) = 104.79, p < .001$). The arousal ($F(1, 50) = 9.162, p < .004$) ratings differed significantly between the round and square avatar. This difference is most likely based on the

Figure 3: The experimental setup.

slight difference of the two avatar's animations. Culture had no significant influence on arousal or valence. Gender had a significant influence on valence ($F(1, 50) = 6.712, p = .013$) and there was a significant interaction effect between culture and gender ($F(1, 50) = 4.398, p = .041$) on valence. Figure 4 shows the means of arousal and valence grouped by the independent variables culture and gender.

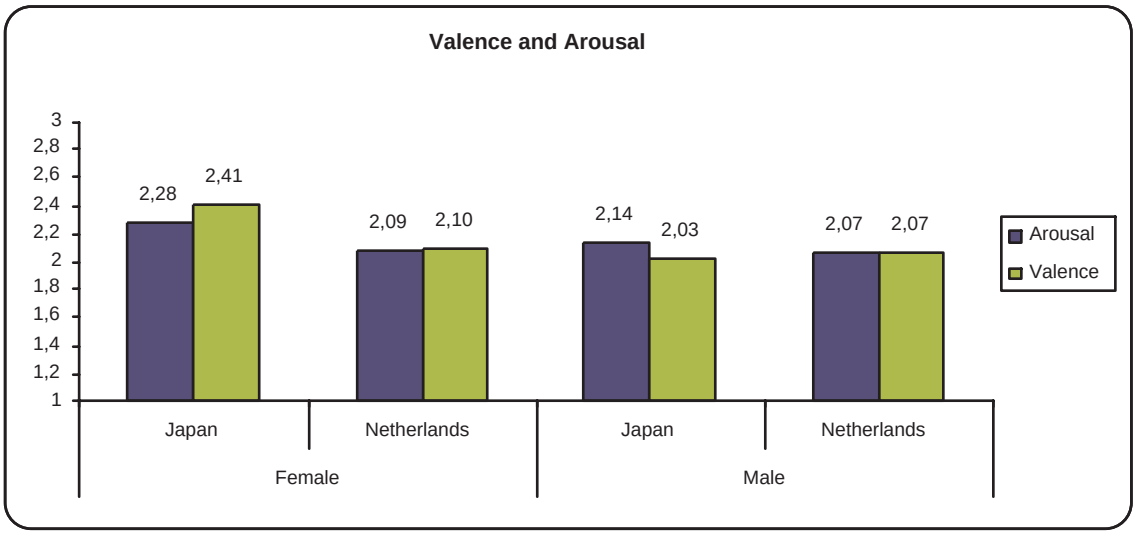


Figure 4: Mean arousal and valence across all animations and avatars

Women rated ($m = 2.136$) the animations significantly higher ($t(53) = -2.624, p = .012$) on valence than men ($m = 1.969$). The Japanese women rated the animations highest ($m = 2.41$) on valence. The animations themselves had the biggest influence on valence and arousal. Certain animations show influence of the culture (see Figure 5 and Figure 6).

Even though no overall influence of culture on the perception of animations could be found, perceptions of certain animations were affected by it. The perception of symbolic and iconic expressions is affected because they carry a specific meaning. Deep bowing, like in animation 26,

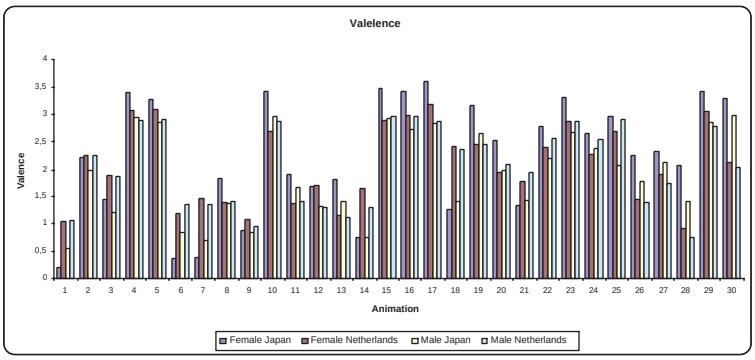


Figure 5: Valence per animation

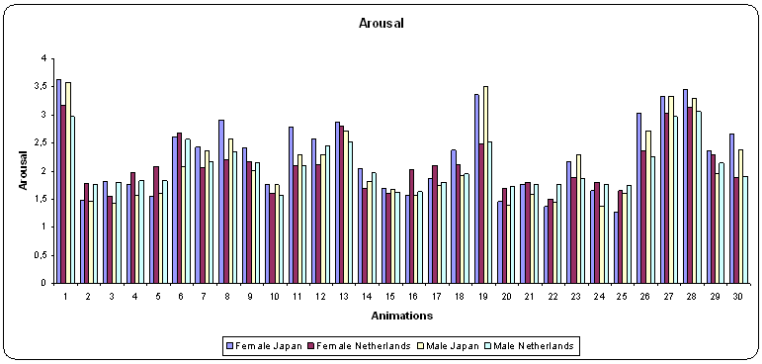


Figure 6: Arousal per animation

is an appropriate form of apologizing in Japan but is uncommon in the Netherlands. However, a deep bow is an extreme expression in Japan and therefore the Japanese participants rated it significantly higher on arousal ($t(53) = 3.813, p < .001$). The same social display rules also explain the difference in arousal for animation 19. Jumping from joy is only performed in Japan under extreme joyful conditions. Therefore Japanese participants rated animation 19 significantly higher on arousal than Dutch participants ($t(53) = 6.5, p < .001$).

Animation 30 is a symbolic expression of "yearning", a Japanese concept for feeling good and being ashamed of it (see Figure 7). This specific concept is not known in the Netherlands and therefore it was rated not as positive by the Dutch as by the Japanese ($t(53) = 5.849, p < .001$).

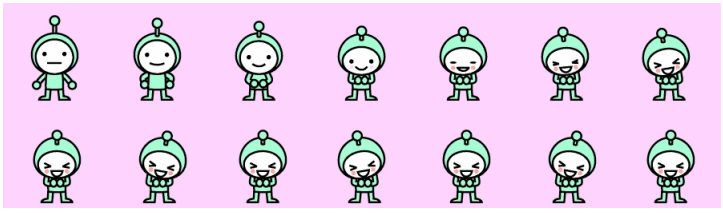


Figure 7: animation 30

Another culture-related difference can be observed in animation 18. Putting oneself forward is inappropriate in Japan and is therefore rated significantly lower on valence by Japanese ($t(53) = 6.563, p < .001$). Dutch participants possibly interpreted it simply as contempt.

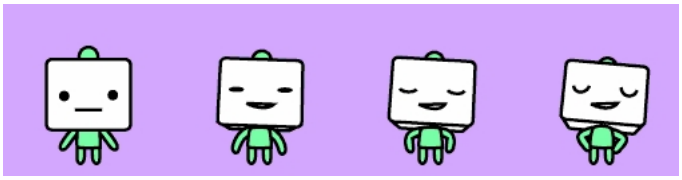


Figure 8: animation 18

4 Conclusions

We conducted a cross-cultural study on how users perceive the expressions of avatars. The animations themselves had the strongest influence. The 30 animations of both avatars were almost identical, but even the small differences resulted in significantly different perceptions. These difference is likely to be even stronger for more diverse avatars, such as a monkey and an elephant. Due to their physiology they will have to express the same meaning in different ways. Developers need to be aware of these differences and evaluate their designs.

Even though we did not find an overall significant influence of culture on the perception of the animations we could identify certain animations that are perceived differently. This differences can be explained by the different display rules in the two cultures. Expressing emotions is only performed in Japan under extreme situations whereas it is more common in the Netherlands. Furthermore certain symbolic expressions refer to concepts that are only known in one culture. The Japanese "yearning" or "bowing" expressions are good examples. The cultural difference concerning bowing is, of course, well documented in most travel guides to Japan.

The gender of the participants had significant influence on the perception of the animations. Particularly Japanese women rated the animations more positive on the valence scale. This important result defines our further research directions. What avatars do users prefer to represent themselves? Should the avatars have a clearly defined gender or is a neutral design better?

References

Damer, B. (1997). Avatars: Exploring and Building Virtual Worlds on the Internet. 1997, Berkeley: Peachpit Press.

Ekman, P. and W.V. Friesen (1971). Constants across cultures in the face and emotion. Personality and Social Psychology, 1971. 17(2): p. 124-129.

Hendrix, J., et al. (2000). A facial repertoire for avatars. in Workshop on Interacting Agents. 2000. Enschede.

Hofstede, G. (1984). Culture's Consequences, International Differences in Work-Related Values. 1984, Beverly Hills: Sage Publications.

Isbister, K., et al. (2000). Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space. in Conference on Human Factors in Computing Systems (CHI2000). 2000. Den Hague: ACM.

Krenn, B., et al. (2004). Life-Like Agents for the Internet: A Cross-Cultural Case Study, in Agent Culture: Designing Human-Agent Interaction in a Multicultural World, P. S. and T. R., Editors. 2004, Lawrence Erlbaum Associates: New York.

O´Neill-Brown, P. (1997) Setting the Stage for the Culturally Adaptive Agent, in Socially Intelligent Agents, Papers from the AAAI Fall Symposium, K. Dautenhahn, Editor. 1997, AAAI Press. p. 93-97.

OSDN (2004). Slashdot 2004.

Rheingold, H. (2002). Smart Mobs: The Next Social Revolution. 2002, New York: Perseus Publishing.

Schlossberg, H. (1952). The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology, 1952. 44(2).

Takahashi, T., C. Bartneck, and Y. Katagiri (2003). Show Me What You Mean Expressive Media for Online Communities. in CHI2003 Workshop on Subtle Expressivity of Characters and Robots. 2003. Fort Lauderdale.

Bridging Communications Across the Digital Divide *

Edwin H. Blake
University of Cape Town
Dept Computer Science, Rondebosch, 7701, South Africa
edwin@cs.uct.ac.za

William D. Tucker
University of the Western Cape
Dept Computer Science, UWC, Bellville, 7535, South Africa
btucker@uwc.ac.za

Abstract

Connecting people across the Digital Divide is as much a social effort as a technological one. We are developing a community-centered approach to learn how interaction techniques can compensate for poor communication across the Digital Divide.
We have incorporated the lessons learnt regarding Social Intelligence Design in an (abstract) device called the SoftBridge. The device allows information to flow from endpoints through adapters (getting converted if necessary), and out to destination endpoints.
Field trials are underway with two communities in South Africa, disadvantaged deaf users and an isolated rural community. First lessons learned show that we have to design user interfaces that allow users to understand and cope with delay (latency) as a necessary consequence of our approach.

Keywords: Community-centered, Digital Divide, Information Society, Multi-Modal, User Interface, User-Centered Design.

1 Information Society and the Digital Divide

One might characterize a developing country as one where there is a great need for better access to resources. Some of these resources are necessarily limited and their distribution is the outcome of a zero-sum game; other resources, and *knowledge* is amongst them, can potentially be distributed to the *have-nots* without taking away from the *haves*. This is a partial answer to the "bread" *versus* "broadband" conundrum.

One of the prime applications of Information Technology (IT) in a developing country is to extend the distribution of scarce knowledge resources. We employ the concept of an *Information Society* to mean the desired outcome of the information revolution sparked by information and communications technology. This technological revolution has universal impact because Information Technology is a Universal Enabling Technology: it acts as an enabler in all aspects of our lives.

The definition of the Information Society makes it clear that Information Technology development depends on the formulation of clear goals for society. Technology cannot be appropriately

*This research was supported by many organizations and people: both fellow researchers and community members. Resources were provided by the NRF, the South African National Research Foundation through its THRIP and Focus Areas programmes, Telkom, Siemens SA and SANPAD, the South Africa – Netherlands Research Programme on Alternatives in Development. The CSIR Information and Communications Technology Section introduced us to village of Tsilitwa and provided the WiFi infrastructure that we used. The NGO "Bridges.Org" acted as a consultant on the applicability of the work. Over the past three years our students have provided many ideas and inspiration, in particular the authors would like to thank John Lewis and Marshini Chetty for developing some of the systems discussed.

applied when what is appropriate is not known. But, whatever the societal goals that are defined, we can assume that IT can provide a cost effective way of reaching *some* of those goals. There are major disparities in the penetration of the Information Society in the developing world. The term *Digital Divide* is to be seen in this light.

"Digital Divide" is the growing gap that exists between those who have access to the Information Society those who are deprived of such access due to cultural bias in the applications and contents, gaps in their education (for example, illiteracy), personal handicap, poor digital infrastructure, or lack of appropriate computer equipment.

The South African Digital Divide grows out of a particular history which is one of great prior division and historical backlogs for large parts of the people: the particular South African version of colonial history. The digital divide is also related to global economic circumstances in which all developing countries find themselves. So it is that we can identify at least two aspects of a Digital Divide:

A. The Global Digital Divide
This is the global disparity between those countries at the forefront of the Information Economy and those developing countries who are striving to enter that system.

B. The Local Digital Divide
This refers to the disparities between the various people in a particular country. Clearly it is much more than simply a digital equipment divide.

1.1 Bridging the Divide

Bridging the Digital Divide is the effort to provide increased access to information and communication to those who have little or none at all. "Communication bridges" involve social dynamics as well as the technological tools that support social interaction. Our community-centered approach has produced innovative systems that provide completely new solutions to the issues that arise in building communication bridges. We support our user communities with new communication systems that are adapted to their requirements.

This must be understood quite literally. We build new computer based artefacts that act as automatic communication bridges between various groups. The basic building block is called a SoftBridge [Lewis et al. (2002, 2003)]. One application is to bridge communications between disadvantaged deaf users by translating from text to speech and back again, Tucker et al. (2003). Another is to provide access to professional medical information for nurses in remote rural clinics, in the face of frequent power and other infrastructure outages, Chetty et al. (2003).

We are developing a methodology to support this design process as well. We have found that sophisticated bridging systems impose delays upon the communication process. Additionally, the unreliable nature of the infrastructure also can result in extended breaks in communication. Thus, compensating for delay is of major importance in building automatic communication bridges over the Digital Divide.

Allowing communications between different groups across the various digital divides requires a number of innovations. The issues we intend addressing are:

1. Changes in the way we design IT based applications and contents,
2. Putting the notion of bridging and translation central to our software systems and building that intelligence into the system, and
3. Altering IT related policy that impedes social development.

We shall consider these in the following sections. In Section 2 we discuss a methodology for designing appropriate IT solutions, then, in Section 3 we present our abstraction that incorporates the Social Intelligence, namely the SoftBridge. In Section 4 we discuss the very first results from deploying our systems in two disadvantaged communities, one urban and the other rural. We

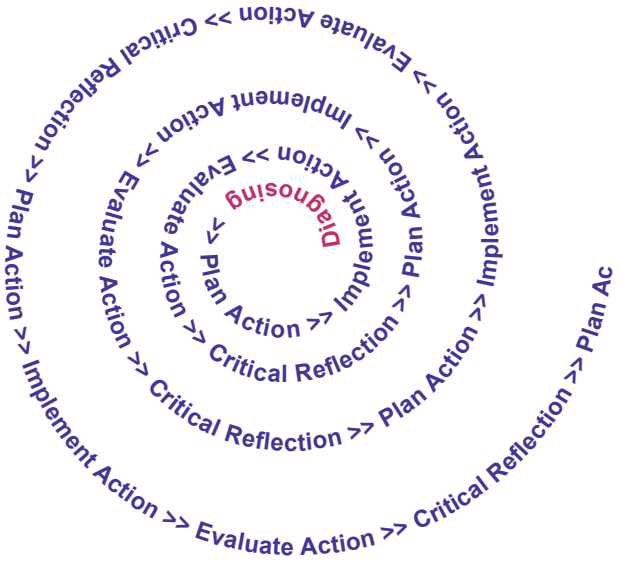


Figure 1: The Cycles of Action Research as a spiral to indicate progress in meeting user needs.

emphasize the need to act on a political level as well in Section 5 and finally we provide a short Conclusion.

2 IT Application and Contents Design Methodology

We believe a user-centered approach is most appropriate to address the design of our systems. We have considered a number of approaches, using various ideas from *Critical Action Research*[Lewin (1948),Stringer (1997),Carr and Kemmis (1991)] design and cyclical software development process based on participatory design and prototype evaluation. Our method retains the cycles of Critical Action Research of diagnosing, planning, implementing the plan, observing results and reflecting on the results. Evaluations of each acting stage form the basis for correctly planning the next step in the process. Each evaluation may lead to modifications in the ultimate goal (see Figure 1).

Critical Action Research has a strong emphasis on the empowerment of groups. It involves facilitating a change in a community through facilitating action. This is done in collaboration with the community members. The flaw in this from our point of view, in terms of design, is that it assumes at least a degree of sophistication of the user community in relation to technological possibilities and an ability of software designers to bridge large cultural and linguistic gaps. This may not be possible.

The cultural gaps can be enormous. The technological requirements exist within a complex web of other needs, relationships and societal obligations. Misinterpretation (on both sides) and unexpected needs are common. It is difficult for IT practitioners to appreciate, for example, how an IT empowerment exercise may threaten power relations in such communities with dangerous consequences for several participants. For many reasons therefore, end-user participation in the process can be problematic. Dray and Siegel (2003) have found a similar need for a wide, and culturally sensitive, view on the software process.

Our approach is therefore to extend the team with advisers and consultants that are drawn in in various stages of the design and development process. We have used the services of non-governmental organizations (NGOs), other researchers already involved with a particular community, and professionals serving the community in other capacities (e.g., doctors). In the initial stages these people form our "human access points" into the community. Rather than focus on empowering the particular individuals who will be the users of the system, we engage a wider

community. We must always be aware of social subtleties. We call this a community-centered approach.

Nishida (2002) has referred to the two aspects of Social Intelligence as firstly the "conventional" one of an individual's capacity to act wisely in accordance with social rules and, secondly, the ability of a system to manage complexity of interaction. We believe this corresponds in our case to the abilities of the designers to discover and operate within the rules of the community (our community-centered method) and secondly to design a system to deal with some of the complexity (our SoftBridge abstraction).

The drawback in terms of experimental results is that such large team design and iterative development processes do not lend themselves to easy documentation of final results. There is no defined group of users that can serve as experimental subjects — we are all co-designers. Therefore we can only strive to combine largely qualitative work with some quantitative empirical results where ever possible.

A measurement system operates alongside the technology development cycle. Starting with an initial baseline, participants are surveyed to determine how they tolerate problems with their communication systems. Further measurements are made as subsequent software modifications are introduced in the field. The software is also instrumented to record actual activity metrics, such as latency. These metrics are correlated with participants' subjective experiences to give a quantitative measure of how well the interaction mechanisms compensate for problems with the bridge. This corresponds to Fujihara (2001) on control conditions: that is using the system with and without the proposed enhancements as a means of measuring effectiveness within a spiral of developments.

3 Automatic Communication Bridges

Our underlying systems level abstraction is that of a generalization of the notion of Quality of Service in networked communication. We call this "Quality of Communication". It refers to the ability of a system to support communication by bridging between different user abilities, sensory and media modalities, infrastructure capabilities and terminal equipment facilities.

In terms of "Social Intelligence" it reflects both the aspect of the social intelligence of users to conduct meaningful exchanges *in spite* of the limitations of technologies ("where there is a will there is a way") and also the aspect of social intelligence of the system where it adapts automatically to the user and system capabilities and bridges between different communication paths (text for deaf users to speech for hearing users, Tucker (2003)).

The bridging is implemented by the *SoftBridge* (see below in Section 3.1 for more details). SoftBridge exists both as an actual implemented Computer Artefact as well as an abstraction for all those intelligent operation that cannot be realized with current technology.

As an actual implemented system [Lewis et al. (2002, 2003)] it can tackle certain bridging operations. It can translate between the format used by the basic South African text-phone for deaf users (Teldem) and Internet based chat. It can break out to the public phone system and it can translate text to speech with a North American accent and has limited success in translating North American accented spoken English into text. There are currently efforts (by other groups) to tackle the eleven official languages of South Africa, including the local South African ("Sef Efriken") version of English. In the meantime the SoftBridge as an abstraction exists and is simulated in practice by a human who acts as the speech to text translator.

What has become clear is that whatever the SoftBridge is, the users will almost certainly always have to deal with *delay*. The processing inherent in the SoftBridge imposes delay. Delay in delivering messages also arises due to poor infrastructure in rural areas. Extended power outages are common place in rural areas. Phone lines can often be down as well (copper wire is stolen). A useful system must be able to switch seamlessly between synchronous and asynchronous modes of communication.

Our first result, namely the necessity of delay, may seem mundane, but it has far reaching implications. Dealing with delay or latency in a Socially Intelligent fashion is therefore a prime problem and a focus of some of our current user based research. The user must retain an appro-

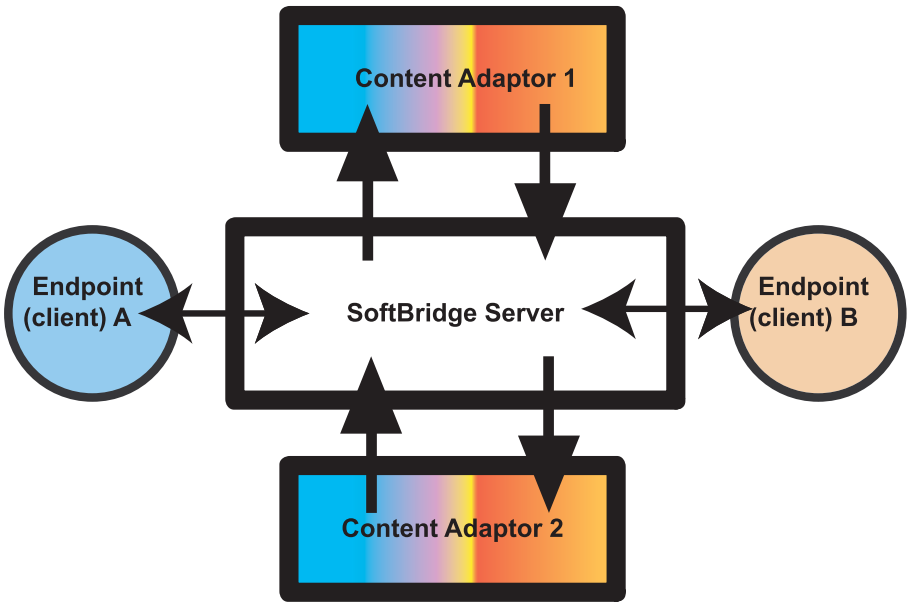


Figure 2: The Data Flow Architecture of the SoftBridge. The diagram shows how the server calls adaptation services to bridge between different endpoint needs.

priate sense of *co-presence* [Zhao (2003)] of the other person that is being communicated with. If the link is down but the person can still be contacted, albeit with delay, that sense of another being there must remain. Alternatively, even if the link is up but the other person has left, that fact has to be communicated.

We are concentrating on adapting communication content and its interfaces to user capabilities. By capabilities, we do not only refer to computer hardware and software, but also the capabilities of the user.

This adaptation process is performed by a system we have called a SoftBridge. It is an embodiment of Social Intelligence Design: both as an abstraction and as a concrete embodiment. The SoftBridge relies on the abstraction of communications mechanisms such as text, voice, video and VR.

A SoftBridge enables us to design and, where practical, build applications that bridge vastly different access equipment (telephones, cellphones, handheld Mobile IP devices, laptops, personal computers, HMDs) to communicate seamlessly using various communications media (text, voice, video, VR) without regard to underlying mechanics of the process.

It also adapts to the user. For instance, if a user is blind, then the system would only deliver audio, or translate text into speech. And if they were deaf, speech would be converted into text. As another example, we can extract human communication from an information and graphics-rich CVE and bridge it to a low end user. Consider a text chat tool inside an immersive Virtual Reality CVE with an IP bridge to a text chat tool running in a web browser. The content (text) and endpoint (HMD and CRT) abilities remain, but the interfaces are bridged. User preferences and profiles can also be taken into account. For instance, if the sender is male, he may prefer that messages from him are read out in a male voice.

3.1 SoftBridge

The architecture is based on the "Data Flow" model of computation (see Figure 2), and consists of the following components:

1. Endpoints.

   Endpoints are the locations where humans interact with the system, and typically consist of devices such as personal computers, cell phones, telephones and wireless handhelds.

2. Content Adapters

   Adapters convert information from a form produced by one endpoint into a form consumable by another endpoint. For instance, if endpoint A produces speech, and endpoint B consumes text, adapter AB will convert speech to text.

3. Control/Location/Billing

   The control function deals with the allocation of adapters and the synchronization of components. It also coordinates the dynamic relocation of adapters as network conditions change. The network conditions are discovered by using Quality of Service (QoS) parameters such as jitter, latency and packet loss. Endpoints register themselves with the Control component when they connect to the system. This allows the Control component to keep track of all currently connected endpoints and simplifies billing.

3.2 An Example Scenario: Bridging Deaf and Hearing

> Andile calls Noxolo with his mobile. She is deaf, so her handheld vibrates to announce the incoming call. Andile's speech is converted to text and relayed to her. She responds in writing that is converted to speech, imitating her voice. The specialized semi-synchronous conversation interface enables them to feel connected despite the delay.

This example illustrates the action of the SoftBridge in translating between the requirements of the two users. Such requirements arise from the combination of personal requirements, the system capabilities at the two ends of the link and the quality of the infrastructure between them. As explained above the SoftBridge will act to link the two users automatically according to these conditions.

A further design issue requiring Social Intelligence is dealing with the inevitable latency issues. There are many possible sources for delay other than transmission latency. For example, the communication could involve bridging from one modality of communication to another, or the communication channel can be inherently unstable.

The user interface can help in dealing with these situations. When a user believes that lack of response means a broken communication link, s/he often gives up. An interface that indicates that a response is on its way could save the situation and encourage a bit of patience. The interface has the potential to provide a social link between the parties even when communication is not being explicitly exchanged. Social software features would address presence, awareness of the other's activities, and persistence of communication threads. These features encourage social connectivity over poor delay situations.

4 Current Status

Field trials are underway with two communities in South Africa.

The Bastion Centre for the Deaf serves a "disadvantaged" community in Cape Town that is marginalized from mainstream communications due to both poverty and hearing disorders. Voice/text relay enables a deaf person to use a "text telephone" to communicate with someone on a normal telephone via an operator with both devices. Because this service is not available in South Africa, we have built an automated voice/text relay system based on web services. Automated Speech Recognition (ASR) weakens the communication bridges by increasing real-time delay via

processing overhead. As noted before, ASR also performs poorly with South African accented English. Preliminary trials have influenced back-end development, Lewis et al. (2003) as well as interfaces targeted for both deaf and hearing users, Tucker et al. (2003).

The second community is located in a remote rural region. Tsilitwa (Eastern Cape, South Africa) has a clinic without a doctor that serves roughly 10,000 people. The Centre for Scientific and Industrial Research (CSIR) has installed a wireless Ethernet network with basic voice and video over IP to allow clinic nurses to communicate with a doctor in a neighboring village. However, the system is rarely used due to frequent power outages. Visits to Tsilitwa, workshops with the CSIR and frequent communication with a local NGO called bridges.org, together have yielded software requirements for a multi-modal store-and-forward system to overcome the power problems. A prototype is now ready for extended testing, Chetty et al. (2003).

4.1 Initial Results from the Deaf-Hearing Bridging Project

As noted before our testing proceeds in cycles. The first cycle of testing of the Deaf telephony project involved a single Deaf user (DU). We conducted three tests varying the input/output modalities of the hearing user (HU). The DU had a standard text in/text out Instant Messaging client. The HU client used the following specific modality combinations: Text and Text-to-Speech (TTS) in/Text out, TTS in/Text out and TTS in/Text and Automatic Speech Recognition (ASR) out. The SoftBridge logged the conversations for subsequent analysis.

The trial showed a largely successful conversation. Success factors included a) a text and computer literate DU who is familiar with research practice, b) using the system to explain the research as we conducted it and c) that the multi-modal bridging capabilities overcame the expected shortcomings of TTS and especially ASR.

It showed:

1. Deaf users use a different grammar (related to South African Sign Language, SASL) and this cannot be automatically corrected for hearing users. This point was emphasized by the deaf user. Thus hearing users have to deal with the poor language skills of deaf users.

2. ASR (Automatic Speech Recognition) is inadequate to the task of recognizing South African English and other South African Official Languages. There are projects elsewhere to address this and in the meantime we will have to mimic ASR by employing a person to provide the service.

3. Presence indicators are needed to show continuation of the conversation when there is no visible activity due to delays.

4.2 Some Preliminary Results from the Rural Tele-Health Project

The Tele-Health software has recently been deployed. This occurred on our third visit to the target community, the first being an orientation visit and the second the presentation of a paper prototype.

In introducing the software we have been struck by the complicated community interactions that determine whether a IT solution will be accepted. For example, the existing power-relations may prevent solutions that help the poorest people in the village.

The essential idea is to allow nurses in outlying villages to communicate with the one or two doctors at the local town hospital. A video conference link is problematic because frequent power failures render it useless for lengthy periods. Due to the shortage of staff at the hospital, tele-medicine is an additional workload for the doctor who is solely responsible for the entire hospital. It is therefore difficult to schedule times when a synchronous tele-medicine consultation can occur.

We decided that combining a store and forward approach with VoIP would resolve the communication problems. Store and forward was explained to participants in terms of voicemail on cellular telephones and everyone was familiar with these. A store and forward approach allows

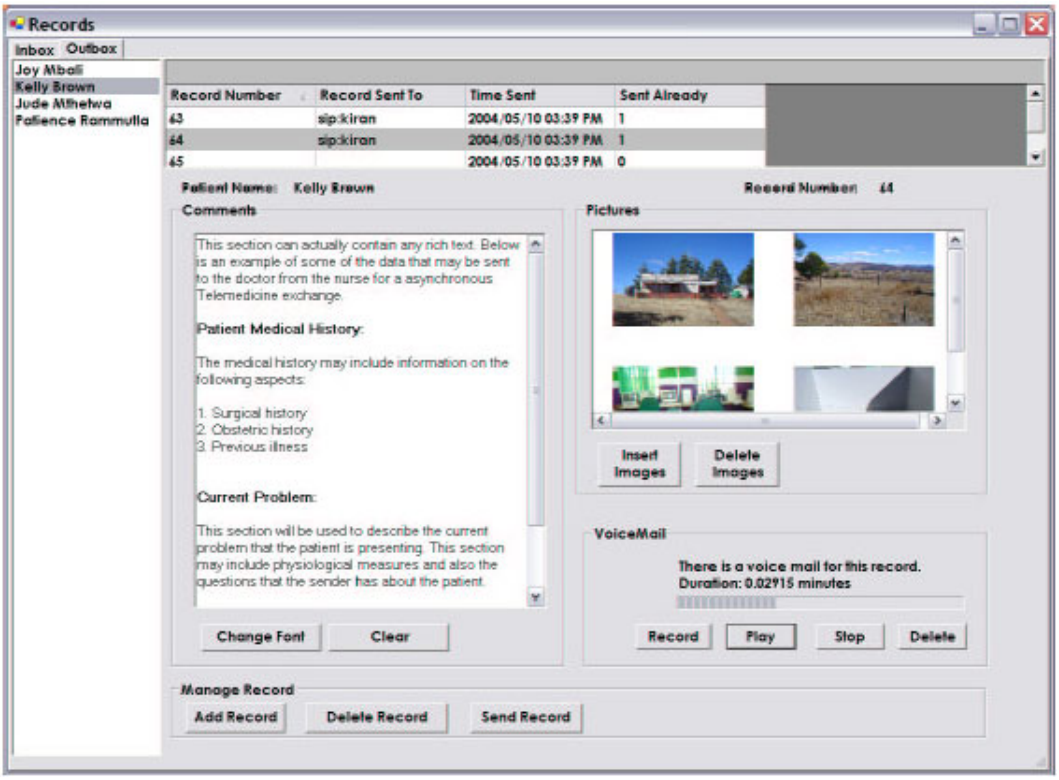


Figure 3: Telehealth Project. This screenshot shows the records dialogue where records are displayed per patient in an Inbox and an Outbox. This screenshot shows a dummy record to illustrate that a record can contain text, images and voicemail. Images may also be viewed in a separate image viewer which allows enlargement.

patient data to be captured at any time and then sent to the recipient site when a network connection becomes available. We decided to introduce laptops since they offer several hours of battery life. This means that data can be captured even during a power failure for as long as the battery lasts.

The software was required to support synchronous voice calls and asynchronous sending of messages between the clinic and the hospital. These messages were to contain text indicating the patients illness and medical history, digital pictures of the patient or particular problem area and voice recordings.

We have now installed the software (see Figure 3) on the new laptops at the hospital and the clinic and connected these to the Wi-Fi network. The initial feedback received from interviews with the doctor and nurses in the area was positive. The doctor felt that our system would allow him to process and reply to messages in his own time without being constantly interrupted during tele-medicine consultations by emergency calls. The nurse at the clinic felt that the system would increase the processing of patients and felt that the digital images and text would help the doctor accurately diagnose problems with a patient.

5 Influencing IT Policy

Social Intelligence with respect to IT is not just a matter for the target communities. As our NGO partners have long known, work on the ground has to be accompanied with work in the upper reaches of government. The Governments of Developing Countries also have to be advised and educated about the implications and potential of IT.

It is necessary therefore to include the influencing of IT policy as one of the essential outcomes of IT research for Development. It is another aspect of a Critical research approach.

In South Africa, for example, the use of Voice over IP is subject to many restrictions. According to SA Telecommunications legislation, VoIP may only be provided by Telkom (the current telecommunications monopoly holder), the long-awaited Second National Operator (SNO) and the Under-Serviced Area Licensees (USALs) [SA Government (1996, 2001)]. USALs are being granted to one Small, Medium and Micro Enterprise (SMME) per geographic area where the less than five percent of the population have access to telecommunications services or facilities. A USAL enables an SMME to provide telecommunications services and use VoIP. In practice this is not effective.

The high cost of network communication and the lack of affordable bandwidth was the most frequently cited constraint on the realization of an Information Society in South Africa. It is setting back education and research. While there have been some very innovative approaches to the problem[1] these cannot be a substitute for changing the policy environment in which we operate.

It is important there fore to generate research results that politicians will understand; and mostly this means ones that will translate into votes. Highly visible applications and good ones that generate good publicity are thus required.

6 Conclusion

We have shown how "Social Intelligence Design" can be applied in a developing country. The aspects of Social Intelligence influence the systems, the users and the researchers. It is especially the IT professionals who have to move from a Software Development Lifecycle based view of systems development, to seeing both the microcosm in which their users operate: the community, and the macrocosm in which everyone exists: the laws of the society.

The implications for IT Artefact development is that one has to develop a methodology that can operate in a very sensitive manner within the culture of the users (our community-centered method) and help those users to deal automatically with the complexities of communicating knowledge (abstracted as the SoftBridge). These are both aspects of Social Intelligence Design.

In the work we have done so far we have shown that, beyond this, when building a device such as the SoftBridge, that dealing with delay is an essential feature. We have argued that this has major implications for user interface design.

References

Carr, W. and Kemmis, S. (1991). *Becoming Critical — Education, Knowledge and Action Research.* The Falmer Press, London and Philadelphia.

Chetty, M., Tucker, W. D., and Blake, E. H. (2003). Using voice over IP to bridge the digital divide — a critical action research approach. In *Proceedings of South African Telecommunications and Networking Application Conference (SATNAC 2003).* SATNAC, Telkom, CD ROM Publication.

Dray, S. and Siegel, D. (2003). Learning from latin america: Methodological lessons from emerging markets. In *Proc. Contextual Invention*, pages 9–18.

Fujihara, N. (2001). How to evaluate social intelligence design. In *Proceedings of Social Intelligence Design 2001.* http://mailer.kc.t.u-tokyo.ac.jp/synsophy/sid2001.

Lewin, K. (1948). *Resolving social conflicts; selected papers on group dynamics.* Harper & Row, New York, 1st edition.

[1] Wizzy: The firm Wizzy (www.wizzy.org.za) has found a unique high bandwidth solution to providing Internet access to deprived schools in townships and rural areas. Transporting the data by road on a high capacity disk! It is a sorry reflection on the skewed cost of bandwidth that one has to resort to such solutions. Wizzy has several other excellent ideas to ensure ease of operation and security and it is able to provide a high-quality solution. By the way, notice the role that latency plays in this solution!

Lewis, J. A., Tucker, W. D., and Blake, E. H. (2002). SoftBridge: An architecture for building IP based bridges over the digital divide. In *Proceedings of South African Telecommunications and Networking Application Conference (SATNAC 2002)*, page 18, Champagne Sports Resort, KwaZulu-Natal. SATNAC, Telkom, CD ROM Publication.

Lewis, J. A., Tucker, W. D., and Blake, E. H. (2003). SoftBridge: A multimodal instant messaging bridging system. In *Proceedings of South African Telecommunications and Networking Application Conference (SATNAC 2003)*. SATNAC, Telkom, CD ROM Publication.

Nishida, T. (2002). Communicative reality for social intelligence design. In *Proceedings of the IEEE Workshop on Knowledge Media Networking (KMN'02)*. IEEE.

SA Government (1996). Telecommunications act no 103 of 1996. Govt Gazette. Available: http://www.doc.gov.za.

SA Government (2001). Telecommunications amendment bill no 65 of 2001. Govt Gazette. Available: http://www.doc.gov.za.

Stringer, E. (1997). *Action Research: A handbook for practitioners*. Sage Publications, USA.

Tucker, W. D. (2003). Social amelioration of bridged communication delay. In *8th European Conference of Computer-supported Cooperative Work*, Helsinki, Finland.

Tucker, W. D., Glaser, M., and Lewis, J. (2003). SoftBridge in action: The first deaf telephony pilot. In *Proceedings of South African Telecommunications and Networking Application Conference (SATNAC 2003)*. SATNAC, Telkom, CD ROM Publication.

Zhao, S. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments*, 12(5):445 – 455. http://mitpress.mit.edu/journals/pdf/pres_12_5_445_0.pdf.

Developing Methods to understand Discourse and Workspace in Distributed Computer Mediated Interaction*

Dr. Renate Fruchter
Project Based Learning Laboratory
CEE Department, Stanford University
fruchter@stanford.edu

Humberto Cavallin, M.S.
Department of Architecture,
College of Environmental Design,
UC Berkeley
hcavalli@uclink4.berkeley.edu

Abstract

This paper presents on-going research towards understanding the discourse and workspace in computer mediated interactions. More specifically we present the development of a series of methods to study non-collocated computer mediated interactions. These methods were developed originally to study interactions involving teams composed of architecture, engineering, and construction management students as part of the AEC Global Teamwork course offered at Stanford University in collaboration with universities worldwide since 1993. The methods stress the value of using ethnographic approaches, particularly the role that both discourse and workspace have in developing the communication processes involved in the interactions. Because of the context of the teamwork course we used as a test bed to develop these methods, we focused on issues regarding the nature of the communication act in building design projects when mediated by computers. We successfully tested these new methods and present initial results.

Keywords: Discourse, Workspace, Distributed Teamwork, Computer Mediated Interaction.

1 Introduction

Recent digital communication technology developments offer groundbreaking advances in geographically distributed collaboration teamwork environments. This real-time distributed collaboration allows people to work together at the same time from different locations on the solution of tasks, as in the case of collocated design processes or management situations (Olson and Olson (2000); Gutwin and Greenberg (1999)) These shared computational workspaces are usually based on hardware and software that enable the sharing of video, audio, data, and software applications. In addition other technologies have been developed based on advanced versions of chat applications, whose original purpose were to enable text-sharing capabilities over the Internet

Inside these high-tech worlds of collaboration, however, the fluency of communication is often affected by the affordances of the interfaces (Chapanis (1975)). Virtual workspaces are often awkward, unnatural, and essentially frustrating compared to collaboration in collocated face-to-face real time settings. Nevertheless, it is not clear what aspects of the work environment create the frustration in these kinds of collaboration conditions.

However the problem of producing research that can enable us to answer these kinds of questions, is not a trivial one. (Cooper et al. (1995); Finn et al. (1997); Gallager et al. (1990); Heath and Luff (1991); Vera et al. (1998)) Conducting research in non-collocated interaction environments involves an unknown landscape that poses several methodological questions:

- How should the collaborative environment be instrumented to collect and store the data?

*This study was sponsored by the Project Based Learning Laboratory (PBL Lab), at Stanford, and the Obayshi Corporation sponsored research project

- How should the data be analyzed and represented so that the results depict a complete scenario?

This paper presents methods we have developed to study non-collocated computer mediated interactions involving teams composed of architecture, engineering, and construction management (AEC) students as part of the AEC Global Teamwork course offered at Stanford University in collaboration with universities worldwide since 1993. The interactions studied have been part of the Project Based Learning Laboratory (PBL Lab) launched at Stanford, an initiative and vision that was in response to the need to improve and broaden the competence of engineering and architecture students regarding the new work settings in the professional practice (Fruchter (1999)). The particular characteristics of computer mediated interactions as communication channels, along with the need to improve the communication process of students involved in this course, formed the test bed foundation to pursue a research program focused on the interactions resulting from the PBL global student teams. These methods have been theoretically grounded on the principle of *Bricks & Bits & Interaction* (BBI, Fruchter (2001)) . This principle defines that these activities occur at the intersection of a conceptual space defined by (1) the design of physical spaces, affordances and limitations of typical preset physical, spatial configurations of workspaces, (2) collaboration technologies, and (3) new ways people behave in communicative events using affordances of IT augmented spaces and content. The BBI principle is based on two hypotheses:

1. If we understanding the relationship between *brick*s, *bits*, and *interaction* we will be able to
   (a) design workspaces that better afford communicative events,
   (b) develop collaboration technologies based on natural idioms that support the activities people perform,
   (c) engage people in rich communicative experiences that enable them to immerse in their activity and forget about the technology that mediates the interaction.
2. Any new information and collaboration technology will require change and rethinking of:
   (a) the design and location of spaces in which people work, learn, and play.
   (b) 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.

This research effort investigates the nature of the communication process of the building design activity when mediated by computers. The research has investigated both the physical setting of the interactions, the workspace, as well as the psychosocial aspects of the interaction, focusing mainly on the nature of the discourse produced in the communication process.

In pursuing this goal, one of the main challenges has been to understand: (1) how to collect the data in order to capture the richness and dynamics of the interactions, and (2) how to formalize it into models and representations that enable a better understanding. This paper presents the methods we have developed to achieve these goals, by recording, representing, visualizing and analyzing data from the interactions, as well as documents the fundamental issues regarding the conceptual communication model we used for these analyses.

2 The Project Based Learning experience

The PBL experience is an ongoing integrated research and curriculum program coordinated by Stanford University involving undergraduate and graduate students from engineering, architecture, and construction management from Stanford, UC Berkeley, and Georgia Tech, Kansas University in the United States, Strathclyde University in Glasgow, UK, Ljibljana Technical University, Slovenia, Bauhaus University, Weimar Germany, ETH Zurich and FHA in Switzerland, TU Delft in Netherland, KTH Stockholm in Sweden, Aoyama Gakuin University, Tokyo and Stanford Japan

Center in Kyoto, Japan. The emphasis of this AEC PBL program is on cross-disciplinary, collaborative, geographically distributed teamwork. AEC teams work on a building project using Internet based communication technologies to connect the participants, enabling them to interact in geographically distributed shared environments. Each A/E/C team has an owner representative whose responsibility is to set the budget, program, and context limitations as well as approve variances to the project. In addition each team has an existing site, and a users program for the building. Other academics and practitioners interested in the research act as mentors representing each of the three professional areas of expertise. Figure 1a illustrates a typical final project.

Students are organized in A/E/C teams, whose members interacted remotely with each other from their home universities in their countries of origin. An example of a typical team would have the architect at Berkeley, the structural engineer at Stanford, and the construction manager at TU Delft, Netherlands, and the apprentice at Stanford Japan Center in Kyoto, Japan. Typically, students use wireless laptops or desktop computers running applications on Microsoft Windows environments (Figure 1b). For 3D CAD designing tasks students use typical applications such as CAD tools (e.g., AutoCAD, 3D Studio Max, Form Z), and simulation tools (e.g., SAP2000, RISA), costing and scheduling tools (e.g., MS Project), as well as custom designed collaboration technologies developed by the PBL Lab research team, at Stanford University (Fruchter et al. (2003a); Fruchter (2003b)). The non-collocated communication tasks are performed by the students use Instant Messaging, NetMeeting exchanging voice, video, and application sharing, as well as sharing design, scheduling, and project cost applications, using Web project group workspaces, and web discussion forums.

Figure 1: (left) Typical final AEC project; (right) Typical videoconference setting in PBL Lab

3 Data Collection

To date, eleven generations of students have participated in the PBL experience. Their learning experience has been the environment for the observations of the teamwork tasks faced through the challenges and opportunities established by the course. The traditional methods of collecting information, e.g. participant observations and notes taking, have proven to be insufficient to account for the complexity of interactions involving computer mediated activities. In order to capture the complexity and richness of the interactions, this research project has used several simultaneous data collection systems – direct observation, video recording, and screen capture.

Direct observation: Observations are a useful resource to complement the media described below. Notes provided by observers about the interactions give first hand information that can escape the capabilities of the mechanical media for recording the information. Observation alone however, poses limitations for producing records of the interactions, i.e., certain important aspects of the events that cannot be easily identified by mere observation. Two other information-tracking techniques have been used in our research work to overcome this limitation:

- *Video recording*: Video is a powerful tool for capturing the interactions occurring in the workspace. It was a major source of information mainly for user interactions in the study of the use of the workspace.
- *Screen capturing*: The information regarding the use of electronic interfaces used by the

participant is also relevant and critical to understand the nature of the tasks being performed by the participants. The images that appear on the monitor of computers used for the interactions by the participants were recorded directly form the CPUs by a video recorder device. This media has shown to be particularly useful for observing the role of desktop's affordability in the interaction.

The use of video to record computer-mediated interaction has been a regular practice in usability studies. There are several ways of recording the information regarding the screen and user, either by keeping the sources of video independent, or by integrating them via the use of video mixers in real time or as post-production activities. In this study, the information was recorded separately and analyzed on an individual basis. The data produced during the seventh generation of PBL totaled more than one hundred hours of videotaped interaction. For the purposes of this study, we used about forty hours of the total recording available. The core of the data was captured on a Video-8 system, using a setting as the one shown on Figure 2.

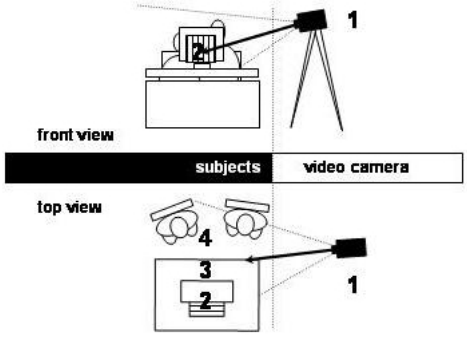


Figure 2: Setting for data collection: (1) Camera, (2) Monitor, (3) Workspace, (4) Participants

The video camera is located in a position that enables the capture of most of the workspace area involved during the interaction, allowing the videotaping of the screen, the input devices, as well as the participants in their interactions. An additional monitor hookup (not shown in the figure) was used for capturing information from the computer screen. The use of video has enabled us to register several relevant aspects of the ongoing interactions of the participants.

4 Data Analysis

In the development of the methodological approach to the study the communicative events in computer mediated non-collocated interaction in building design, one of the aspects considered to be important to stress was the complex nature of the processes involving both communication and the use of the workspace. Traditionally, evaluation of computer mediated interactions have been studied from perspectives that prioritize the quantitative actions of the interaction, i.e. time employed in the solving of tasks, shortcomings of the qualitative aspects of the interaction in terms of the content and structure of the communication processes, and use of the workspace. Moreover, either a purely qualitative or a purely quantitative approach will result in giving a partial account of the activities under study. Considering this, our approach is grounded in a methodology that seeks to consolidate both perspectives, bringing a more complete description of the processes involved in computer mediated teamwork.

The perspective presented in this paper includes two elements traditionally observed by usability studies: the psychosocial processes and the physical objects involved in the interaction. Many of the recent approaches to the study of the use of human computer interaction have focused on these two aspects for producing explanations about the nature of the activities involved in the interactions and the role of computers as mediators. Some of the studies particularly stress the role of the social components in the interaction's diverse theoretical frameworks, as in the case of the activity theory developed by Vygotsky and Leont'ev (Bodker (1997); Fruchter (2004)).

This approach used in the social sciences and education has been applied in the field of HCI to the study of computer mediated task solving.

For the purposes of this research work, the different elements involved in the interactions have been organized into two major categories:

- Psychosocial processes: Diverse processes involving both aspects of the individual and the group psychology, take place in a non collocated interaction. The main resource selected as a way for accessing these processes has been the discourse produced by the participants in the interactions. The verbal activity of the interaction has usually been the object of study of those concerned with the use of protocol analysis as a means of evaluating computer interfaces. Traditional protocol analysis however, seeks information embedded in the spoken discourse by means of quantitative processing of the verbal data, e.g. repetition of words and topics, without necessarily considering any of the other aspects included in the interaction, e.g. the limitations placed by the tools used in the communicational process. This study of the discourse produced in the interaction, takes into account both the level of the structural aspects of the discourse as well of those concerning meanings.
- The nature of the workspace: In addition to the analysis of what is said, this study shows the relevance of extending the analysis beyond the verbal into the characteristics and affordances of the workspace. It is important to include the nature of the relationship between what is said in the context of how that discourse is staged in the physical space as well as the equipments used for communicating.

Finally, it is relevant to consider the psycho-sociology of non-collocated communication both in the asynchronous mode, as in the case of email lists and news groups, as well as synchronous, in the case of chat rooms, MUDs, and avatars. The social nature of the work developed in PBL Lab emphasized the understanding of the different actors in the non-collocated interaction as members of a particular community or group. The processes occurring at this psychosocial level can explain aspects of the interaction such as the developing of trust, emotional bonding, and style of cooperation.

4.1 Identifying the Psychosocial Processes

The discourses produced during the interactions were analyzed with the premise that they were more than a mere exchange of verbal information. In the context of the methodology developed for this study, the discourse is analyzed as a constituent element of a bigger structure of events that include others aspects, such as the use of tools and their role on the production of such a discourse. In order to produce a representation of the situation that would keep as much of the richness of the events of the interaction as possible, schemas and procedures were developed for collecting, coding, and representing the analyzed situations.

Two main categories were considered when coding the data:

1. Discourse data: Discourse information corresponding to the different verbal exchanges occurred during meetings. They were transcribed coding not only the purely verbal aspect but also relevant non-verbal aspects, as gestures, voice intonation, interjections, expressions, and idioms as the speakers used them in the discourse. Interruptions, delays, and pauses also were noted as previously illustrated. All these elements were particularly useful in determining possible communication intentions and rhetoric strategies by the speakers.
2. Behavioral information: Not all the relevant information produced during the interactions was, as it has been pointed out before, verbal. An important part of the interactions corresponded to aspects that involved body language and the use of the hardware available during the interaction. Some of the issues that were coded in our transcripts regarding these aspects were:

   a. Description of the gaze of the speakers: The direction where the gaze of the speaker was pointed provided information about the aspects of the interaction on which the speakers and receivers focused.

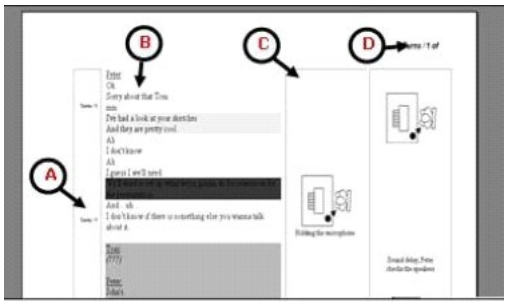


Figure 3: A typical transcript page included three fields: Time stamp (A), Verbal discourse transcript (B), and General Cues (C). Header (D) containing the information pertaining the particular interaction transcribed

b. Notes about the use of software: Describing aspects such as applications used during the meetings, movement of the mouse, location and size of the windows opened, etc.

c. Notes about the use of the hardware: this aspect is very important because it gave clues about the use of the computer and other equipment. Some of the aspects that were tracked were the use of microphone, video camera, and pointing device(s).

d. Movements of the participants in space: This aspect provided information about the use of the available resource in the environment related with space and hardware. We formalized these observations in diagrammatic representations using an upper view of the participants.

A typical transcript page is illustrated in Figure 3. The page is divided into four fields

1. Field (A) was used to mark the time elapsed during the interaction. Timing was extremely valuable information, both when going back to the source of the text and when matching different components of the same interaction. From the experience of this study it is recommended that the time intervals be no longer than 30 seconds because it makes the identification and retrieval of any particular event in the discourse during the analysis process more accessible.

2. Field (B) contains the transcript of the verbal discourse. Each speech turn in the interaction was identified with the name of the speaker.

3. Field (C) provides general cues, such as the description of the gaze of the speakers, notes about the use of objects, notes about the movements of the participants in the space and notes about the use of software and hardware. The use of space was represented by a plan view of the speaker and their position in relation to the computer.

4. Field (D) indicates the name given to the meeting when the tape was identified during the recording, the number of the page, and the number of pages of the transcription.

In analyzing the discourse, the smallest unit of communication for this level of analysis was the turn defined by each intervention produced by any of the speakers in the context of an interaction. This indicates that the conversation takes place through alternate turns among the participants, even though the turns can be incomplete utterances. The turns were grouped and structured into larger units, creating three different levels of aggregation inside the discourse's structure.

1. Topics: the topics correspond to those identifiable themes raised by the speakers during the conversation. Different turns can share a same topic.

2. Episodes: episodes are series of turns that share some specific functional content in the context of the discourse. These turns in the episode can belong to different kids of topics.

3. Protocols: protocols point out the existence of patterns in the communication between the participants that happen in the inner structure of the Episodes. A protocol is shaped by a particular series of turns, which produce structures of verbal and/or behavioral actions that can be identified as having a particular purpose in the context of the interaction.

The following example illustrates this distinction. Consider the following text extracted from one of the interactions analyzed:

```
A: Ok. That sounds good. So, I'm meeting with R at 8:30
B: Ok... Ok... I... and... Ere... Do you... I don't think... I don't know...
   I don't think we should make it a videoconference, cause I don't have
   too much to say of it. What do you think?
A: I agree. I sent an e-mail to R already about that.
B: Ok... That sounds great... So... Uh... About tomorrow... We have to speak
   about 1/2 an hour for the final presentation.
```

The participants produced four turns in this segment two turns each. During these four turns three topics were presented: meeting with R, the need for having a videoconference, and the presentation for the day after. These sets of turns belong to an episode that involves a discussion about the planning of a series of meetings. In this example there is a recurrent strategy employed by B during the interaction consisting of a series of interruptions and interjections that make the transition between the topics. This strategy is called a correction/restart (Goodwin (1981)) and in this case, because of its recurrence as a strategy for taking over and introducing changes in the topics of the conversation, we consider it as a protocol used by B as a control strategy during the interaction.

Once the identification of the different levels of aggregation of the discourse were identified, two kinds of analysis were applied. The first was a variation of the Linkography, a technique developed by Goldschmidt (1998). This technique enables to graphically represent the relationship among the different topics present in the discourse. Linkography is useful to identify characteristics of the verbal interaction as the different topics enclosed by each episode, the connections between topics, and the recurrence of them. Figure 4 shows one of the Linkography produced. It is possible to identify in the graphic the recurrence of the topics, represented by the several triangles; the bigger the triangle, the farther the appearance of a topic is from the last time it was mentioned in the interaction.

In order to make this information explicit we augmented Goldschmidt's Linkography method by color-coding the triangles. Use of colors made it possible to identify important parts of the topics, for example, those that are more frequent interactions (small series of pyramids in Figure 4) are the topics that constitute he core of the discussion. A second graphical analysis was applied to the data, based on the measured distance between the uses of particular topics in the discussion (Figure 5). In order to represent the distance, the number of topics that appear in between the appearance and reappearance of a particular topic was measured. The higher the value represented in the graphic, the larger the amount of topics that separate the appearance and reappearance of a particular topic. In the graphic, the horizontal axis enumerates the appearance of the topics, and the vertical axis corresponds to the value of the distance between the topics, that is the amount that different topics that appeared since the last time the same topic was mentioned in the discourse. It is possible to identify in the graphic four well distinguished phases by using the same example we previously selected to show the use of Linkography, (Figure 5). These phases correspond to what we defined previously as episodes in the interaction.

A third method used to represent the verbal interactions and the different episodes contained in them was the Bar Graph (Figure 6). These bar graphs consisted of horizontal bands in which the horizontal axis represents the temporal duration of the episodes represented in a proportional scale. This temporal reference is important to establish ratios between the weights of the different episodes in the context of the whole conversation. The readability of these graphics is very straightforward, and enables easy identification of issues as recurrences and proportional relation

in the different episodes.

Bar graphs also illustrate a third level of aggregation as mentioned before that is the identification of the communication protocols present in the different episodes. An example is illustrated in the strips pointed out by the arrows in Figure 6, that represent the protocols used for producing the transition between episodes.

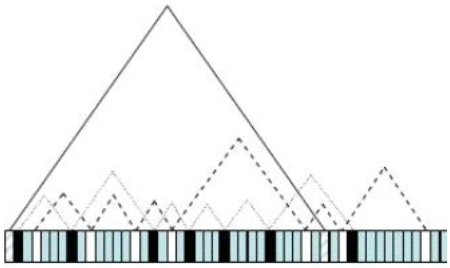


Figure 4: Linkography representation of a meeting. Boxes on the bottom correspond to different Topics. The recurrence of the Topics is represented by several triangles on top. A different line type represents each topic.

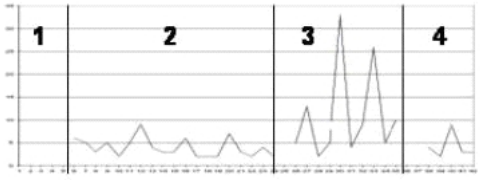


Figure 5: Distance between topics appearance in a non-collocated meeting. It is easy to identify the presence of four different phases: (1) greetings, (2) topics and discussion (3) conclusion, and (4)farewell

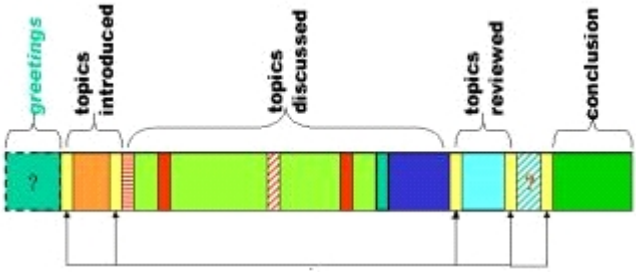


Figure 6: Bar chart of the episodes during an interaction. The horizontal axis represents the time of the Episodes. Each segment represents different episodes during the interaction. In this case, i.e., the five episodes marked by arrows represent episodes involving transition between the different major phases of the interaction.

4.2 The Workspace

The workspace is a central component of our analysis of computer mediated communicative events in global teamwork. From the observations made of the interactions, one important aspect in understanding the behavioral patterns of the participants was the study of the workspace in which interactions took place.

Most of the time, the environment was pre-configured, fixed and constrained to the desktop and its immediate surrounding areas, creating a very restricted space for the equipment required for the interaction. Other relevant aspects of the workspace were the affordances of the equipment used, which co-determined the way in which the participants used the space. In order to represent the use of the space, a graphic code was developed based on a to scale floor plan representing of the participants and equipments used in the interaction. The use of this graphic representation, particularly two aspects have shown to be very relevant in the way in which people use their workspace:

(a) the locations of the monitor and

(b) the placement of the digital video camera used for the videoconferences.

(c) Both monitor and video camera placement define preferred locations for participants, narrowing the possibilities for using a certain area in the working space.

5 Applying the Data Analysis Methods

The following examples show three applications of the previously described methods. On these cases, it is possible to see how the combined approach of analysis involving discourse, interaction, and workspace, has enabled us to identify processes and dimensions that would have been almost impossible to determine by using only one of the listed categories of analysis.

5.1 Visibility

The Figure 7 illustrates the movement pattern for three participants – one in Berkeley and two at Stanford - during a video mediated interaction. Both non-collocated components – Berkeley and Stanford - of the conference are represented back to back. It is possible to see that the movements of the participant on the left side are more confined to a particular area than those of the two people that are sharing the cone on the other side. The special movement analysis over time provides an insight into the use, affordances, and limitations of computer augmented workspaces.

The analysis of the scenario illustrated in Figure 7 describes the environmental aspects present in the interaction. An example of it is the analysis of the participants' movements in the space as it was described in the methods for transcript. In this example, the occurrence of the movements was analyzed in time, correlating it to the verbal interaction, and physical in relation to the computer's location. By crossing the information about episodes and movements in space, we were able to see, i.e., that there is an increment in the physical movements of the participants once a second participant arrives. We were also able to see the effect of the location of the video in the use of space camera watching carefully the disposition of the participants in the space. In the cases studied, the presence of the camera forces a diagonal disposition of the participants during most of the interaction

Figure 8 illustrates the total amount of movement during the communicative event shown in Figure 7 as a result of overlapping the vignettes. From this analysis the overall amount of movement occurring in the workspace emerges.

From the observations we made of the interactions, one important aspect in understanding part of the behavioral patterns of the participants was the study of the workspace used in the interactions. It was reduced in general to the area surrounding the computer, creating a very restricted space for the equipments required for the interaction. The affordances of the equipments used also determined the way in which the participants used the space. Particularly, two aspects have shown to be very relevant in the way in which people use the workspace: the locations of the Monitor and the Video camera.

Both Monitors and Video cameras define preferred locations for participants that narrow the possibilities for using a certain area in the working space. When analyzing the movements of the participants in relation to the location of the equipment, we can see that the movements are restricted to a triangular area, that we have called Cone of Interaction (COI). Four major areas of the COI have been identified through our studies (see Figure 9):

1. Command area (A): Is the area in which the person that leads the interaction is located. The position is determined most likely by the use of the input device, and the fixed hardware configuration is usually geared left-handed users.
2. Secondary area (B): this is the area occupied by default by the other person collocated participants involved in the interaction.
3. Pointing device area (p) is in this case the mouse.
4. Microphone area (m) is the location where the microphone is placed.

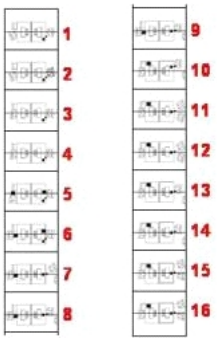


Figure 7: Series of frames showing movement

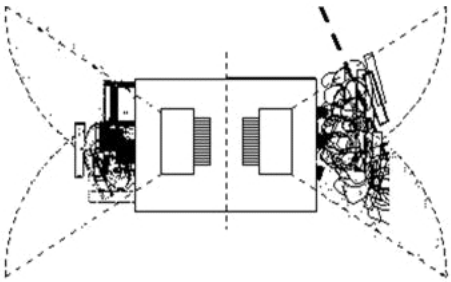


Figure 8: Representation of total amount of movement

The existence of the COI does not imply that participants necessarily use the whole area provided by the cone during an interaction.

Some other aspects have to be considered in relation to the COI. On one hand, the overlapping of functional areas created by the video camera and the monitor. This overlapping creates three zones in the specific workspace (see Figure 10):

- Zone (1) defines the area in which the user of the computer can be aware of the shared digital content on the monitor screen; as well as be visible to the remote participant(s) as he/she can be captured by the lens of the video camera;
- Zone (2) is most problematic of all, because when being in this area the user of the system can be aware of the content on the computer screen, but at the same time be out of the camera range without noticing it, i.e., not visible to the remote participant, therefore creating a visual contact failure during the communicative event.
- Zone (3) is the area in which no visibility or awareness is afforded, for both the user of the computer and for the receiver of the image captured by the video camera.

The COI helps identify which areas of the workspace contribute to a false sense of awareness of the participant in the videoconference interaction, by creating the wrong belief that being in the visibility range of the screen, the actions performed will be transmitted to the non-collocated participant.

5.2 Discourse and Workspace Analysis

The three representations described in the previous sections show partial information regarding the interaction, i.e. structure of the discourse, movement of the participants in space, use of the workspace. When this partial information is integrated, a richer representation of the interaction emerges. The following example shows some of the conclusions derived from a complex representation of the interaction produced by crossing information resulted from the three interaction analysis perspectives some previously described. For instance, a data representation was produced

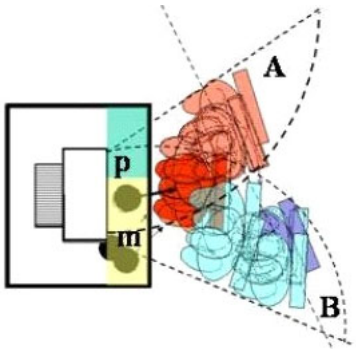


Figure 9: Four major areas in the Cone of Interaction (COI). Command area (A), is the area in which the person that leads the interaction is located, Secondary area (B), this is the area occupied by default by the other participant(s) involved in the interaction, Pointing device area (p), and Microphone area (m)

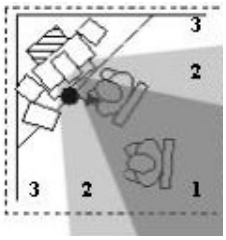


Figure 10: Plan view showing the different areas of visibility and awareness

based on the analysis of the participants' movements in the space as described in the methods for transcribing in a previous section of this paper.

In this example, we perform a temporal analysis that links movement and discourse in time. The amount of movements performed by the participants was analyzed in relation to occurrence in time. By correlating movements of the participants to the verbal discourse during the interaction and to the workspace, i.e., the participants' relation to the computer's location, a single graphic representation can depict information about the structure of the episodes and amount and nature of the movements of the participants in space, as shown on Figure 11.

The colored bar on top of Figure 11 shows a series of interactions represented in a proportional time scale, as was shown on the meta-structure of the episodes. The lower part of Figure 11 shows plan views of the physical disposition in space of the participants during the interaction as it was shown in Figure 7. As it was pointed out earlier, a second person joins the meeting at a certain point in time, and after that we can see that the amount of movement in the space produced by the participants on the two people side increases compared to the instance when only one person was present. There are eight different spatial situations in approximately one third of the interaction time, in contrast of having only two different spatial situations during the first two thirds of the interaction. The arrival of the second person represents an increment of almost eight times in the amount of movement.

It seems that the amount of movement of the participants and the contents of the discourse have no correlation, because during more than half of the interaction the movement of the participants was reduced to a minimum. Moreover, the analysis of the series of vignettes represented on Figure 7 illustrates that the amount of movement in the one-person side stays almost without any change of movement during the whole interaction.

Correlating the information represented in this graphic with the information provided by the analysis of the COI for the depicted situation, it is possible to correlate the nature of the affordance of both the screen and the workspace to the behavior of the participants. After the arrival of the

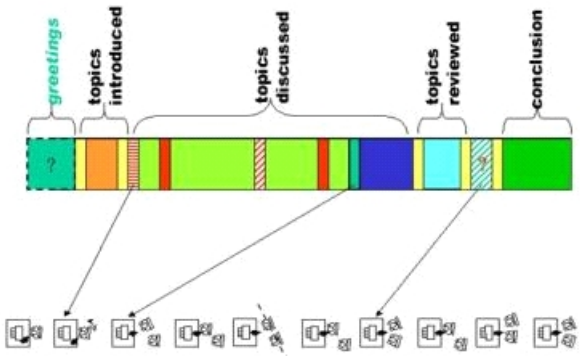


Figure 11: Series of interactions represented in a proportional time scale correlated with top views of the workspace. Second person joins at a certain point during the interaction, and after that the amount of movement in the space produced by the two participants increases compared to the situation in which only one person was present.

second person, they have to alternate their location in front of the screen in order to be able to address the receiver on the other side of the connection. Also, they alternate their turns on using both microphone and mouse, what results in producing the interaction we can see in the vignettes.

By comparing this physical situation with the nature of the verbal discourse developed during the interaction, it is possible to observe that the workspace creates a condition of unevenness for the couple using the workspace, in which one out of the two exert the control over the communication channel, by means of the position in front of the screen and camera, proximity to the microphone, and use of the mouse. This creates a situation of disadvantage for the other participant, which is reflected by the amount of interaction that is allowed to produce in relation to the other participants in the meeting.

5.3 Awareness

An example of the impact of wrong sense of awareness during non-collocated interactions, is an action that we have found often happens when participants are describing information that involves graphical information. We have called it "Faked pointing" (**Fp**), and it points out to a communication failure situation, in which wrong assumptions about the contents that are transmitted with the interface can lead to misunderstanding and delays in the communicational process. Due to its particular characteristics we have decided to analyze it. Consider the following segment extracted from an interaction (Figure 12):

```
Let me tell you this...Look this up
If I don't give you a scale very soon
I will give you the size of one of those rooms
So you can work out the scale off of one of those ratios
But, I think that this is printed at a scale of 1/32 of an inch in a foot
But yeah, the the the...long side is a 116 feet...Um...or actually
um... it's actually 116 feet those little offices uh...
I'm going to take control for just a second. Uh...ok, right here where
the mouse is...These small offices
These are the student offices and uh...they are 10 feet by 6 feet
```

The images in Figure 12 show the gestures of the speaker while describing the graphics on the screen. These gestures, that would be meaningful in a face-to-face interaction, are completely lost in the computer-mediated communication because they fall out of the range of vision of the video camera. What the speaker meant by *"long side"* and *"those little offices"* most probably was unknown by the receiver of the information. The speaker himself notices after a certain point that

Figure 12: Fake pointing examples during a videoconference session

the information he is aiming to transmit has not been conveyed by his gestures and explanations, and decides at the end of the segment to take control of the mouse and use it as a pointer for indicating the object of his description. The use of body gestures and its use for conveying the discourse are drastically constrained by the affordances of the video devices in use, and the lack of awareness of this fact by the speaker can lead to important losses in the communication process.

The role of visibility and awareness in videoconference-mediated geographically distributed collaborative, communicative events, seems to be relevant in situations in which the participants are involved in dialogues that imply the exchange of visual information or social issues. The social value that the visual conveyed by the use of video adds to the interaction and resolution of tasks is highly relevant (Veinott and Xiaolan (1999)).

A better understanding of the role of visibility and awareness in videoconference-mediated teamwork by studying the practices that are involved in the communication process are useful in identifying situations in which the lack of awareness can lead to miscommunication situations. Figure 13 shows how by reconfiguring the workspace, e.g., making small changes in the location of the hardware, part of the problem of the zones of visibility and awareness can be solved. In this case, the camera was located slightly behind the monitor, making the angle of capture and that of visibility to overlap as much as possible.

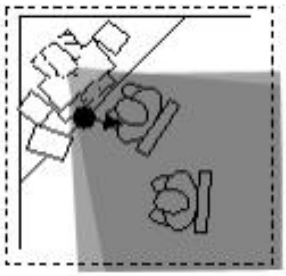


Figure 13: Re-configuring the workspace to address visibility and awareness affordances

Another factor to consider is the physical distribution of the workspace. It has to provide space not only for the computer (CPU, monitor, etc.) but also for activities as writing or use of other equipment, all inside of the COI.

Using interfaces that provide enhanced visibility of the participants, as in the case of Ishii's Clear Board, could be a better solution to the communication problems we have pointed out in this paper. However, interfaces as the one mentioned have been designed for the participation of two people. In order to develop hardware for multiple participants' interactions, it will be required to define new and creative hardware options that will enable to tailor the interfaces to the demands of the activities and natural idioms of users.

6 Conclusions

Computer-mediated interaction between humans involves processes that lie beyond the scope of studying only the technology used in the exchange of data. The complexity of the processes

belongs to the realm of complex communication events, in which the computer is only part of a much larger system involved in the interaction. Once the study of computer-mediated interactions is taken out of the computer realm, it is clear to what extent other aspects different than the tool itself are involved in the processes of communication. The interrelation between these aspects defines the nature of the interactions, and affects both *the way in which things are said* and *what is said.*

In order to reconstruct this complexity through research inquiry it is necessary to collect, analyze, and represent information from different domains e.g. psychology, sociology, and human factors. A major aspect in the communication process analyzed was the verbal discourse produced by the participants. As we illustrated, the words themselves are insufficient to define what occurs during the interaction. In order to better understand such discourse a parallel study of the participants' behaviors was a relevant way to access deeper layers of meaning in the communication event. Behavior in some cases becomes a predictor of the discourse, and in others a negation of the verbal communication itself.

Another important aspect shown by the study is the role that the physical setting plays in the development of the interaction. The processes involved in any communicative action implied much more than what was said in the verbal discourse. Particularly in the case of computer-mediated communication, this study shows how the natures of the interfaces used as well as the affordances provided by the workspaces, define the way in which the interactions occur. The affordabilities defined by the workspace highly determine the way in which the communication process occurs. The spatial quality of the workspace defines the way in which both equipment and people can be allocated within the setting in relationship to computer devices, but also the computer hardware defines the way in which people allocate and use the tools. It becomes an intricate pattern in which equipment, space, and subjects intertwine to define the setting and framework the communication event that defines the interaction. This paper presented methods to extract key elements to analyze the rich nature of the interactions, the processes involved, and the environment in which non-collocated collaboration can take place. Pursuing further research on this methodological direction will continue to require a broad set of fields of knowledge in order to understand and improve computer-mediated communication. As Carroll (2001) emphasizes, future research to augment software quality and performance, needs to continue to focus on areas such as usability and human interaction as core areas of study. Furthermore, we stress the need to concurrently study elements related to the architecture of the workplaces, and the actions that constrain and connect the immediate environment, i.e., *bricks*, to the computer interface and hardware, i.e., *bits* (Fruchter (2001)) as a holistic approach to contextualize future research.

References

Bodker, S. (1997). Applying Activity Theory to Video Analysis: How to Make Sense of Video Data in Human-Computer Interaction. In B. Nordi "Context and Consciousness", Cambridge: MIT Press.

Carroll, J. M. (2001). Human Computer Interaction in the New Millennium. New York: Addison Wesley.

Chapanis, A. (1975). Interactive human communication. Scientific American, 232, P. 36-42

Cooper, G., Hine, Ch., Rachel, J., and Woolgar, S. (1995). Ethnography and human-computer interaction. In Peter J. Thomas (Ed.) The social and interactional dimensions of human-computer interfaces. Cambridge: Press Syndicate and the University of Cambridge.

Finn, K.E., Sellen, A.J. and Wilbur, S.B. (Eds.) (1997). Video-mediated communication, Hillsdale, NJ: Lawrence Erlbaum Associates.

Fruchter, R. (1999). "Architecture/Engineering/Construction Teamwork: A Collaborative Design and Learning Space," *Journal of Computing in Civil Engineering*, October 1999, Vol 13 No.4, pp 261-270.

Fruchter, R. (2001). "Bricks & Bits & Interaction," Special Issue on *"Exploring New Frontiers on Artificial Intelligence,"* Eds. T. Terano, T. Nishida, Akira Namatame, Yukio Ohsawa, Shusaku Tsumoto, and Takashi Washio, *in Lecture Notes on Artificial Intelligence* (LNAI) 2253, Springer Verlag, December 2001, 35-42.

Fruchter, R., Demian, P., Yin, Z., Luth, G. (2003a). "Turning A/E/C Knowledge into Working Knowledge," *Proc. of Computing in Civil Engineering Conference in conjunctions with ASCE National Convention*, ed. Ian Flood, Nashville.

Fruchter, R. (2003b). "Innovation in Engaging Learning and Global Teamwork Experiences," *Proc. of Computing in Civil Engineering Conference in conjunctions with ASCE National Convention*, ed. Ian Flood, Nashville, 2003b

Fruchter, R. (2004). "Degrees of Engagement," *International Journal of AI & Society*, accepted for publication, 2004.

Gallager, J., Kraut, R.E. and Egido, C. (Eds.) (1990). Intellectual Teamwork: the social and technological foundations of cooperative work, Hillsdale, N.J.: Lawrence Erlbaum Associates

Goldschmidt, G. (1998). "Contents and Structure in Design Reasoning". Design Issues, 14 (3), pp. 85-100.

Goodwin, Ch. (1981). Conversational Organization: interaction between speakers and hearers, New York: Academic Press

Gutwin, C. and Greenberg, S. (1999). "The effects of workspace awareness support on the usability of real-time distributed groupware". *ACM Transactions on Computer-Human Interaction*, Volume 6, No. 3 (Sep. 1999), P. 243 - 281

Heath, Ch. And Luff, P. (1991). Disembodied conduct: communication through video in a multimedia office environment. In *Human factors in computing systems conference proceedings on Reaching through technology*, P. 99-103

Olson G.M. and Olson J.S. (2000). "Distance Matters". Human Computer Interaction, Volume 15 (2000). P. 139-178.

Veinott, E. and Xiaolan, F. (1999). Video helps remote work: speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of the Conference on Computer Human Interaction* (CHI'99), Pittsburgh, PA, P. 302-309

Vera, H.A., Thomas, K., West, R.L. and Lai, S. (1998). Expertise, Collaboration and Bandwidth. In *Proceedings of the Conference on Computer Human Interaction* (CHI'98), Los Angles, CA, P. 503-510

Wertsch, J. V. (1991). Voices of the mind. Cambridge: Harvard University Press. `http://www.media.mit.edu/~ishii/CB.html`

Social Intelligence, Cooperative Behaviour, and Collective Action

Satinder P. Gill

Abstract

Knowledge as episteme resides in linguistic and empirical validity, expressed in rules, logic, and conventions. Knowledge as performance, or know-how, resides in experiencing, beliefs, values, emotions, attitudes, and responses. Understanding knowledge as skilled performance in cooperative action, is core to the discussion in this paper. Skilled cooperative action means being able to understand the communicative situation and know how and when to respond appropriately for the purpose at hand. This skill is of the performance of knowledge in co-action (Gill and Borchers, 2003), situated in the formation of legitimate peripheral participation (Lave and Wenger, 1991), and is a form of social intelligence for sustainable interaction. Social Intelligence, here, denotes the ability of actors and agents to learn and to solve problems as a function of social structure and to manage their relationships with each other. A challenge is to create mediating interfaces that can facilitate social intelligence and become invisible to us, in co-action.

keywords: coordinated autonomy, collective action, knowledge in co-action, parallel coordinated moves, social intelligence, overhearing, legitimate peripheral participation.

1 Social Intelligence and Collective Action

Social Intelligence denotes the ability of actors and agents to learn and to solve problems as a function of social structure and to manage their relationships with each other (Nishida, 2003). 'Collective action' implies this social structure (groups, teams, etc.) and 'purpose' (an idea of what counts as accomplishment, etc.)[1] essential to social intelligence for the functioning of any organisation. Collective action is shaped by the movement between the individual and social situation, and we achieve this through our understanding of the performance of representations of the tacit dimension in our communication, such as gestures, non-verbal cues, speech, and pauses, as well as artifacts of practice, including technology, evident in how we perform with them. This understanding mediates the constant transition between individual and social states. Expertise in communication, or being a skilled performer of knowledge in a team, involves this ability to move effectively between 'overhearing'[2] and actively participating. It is about social intelligence.

2 Knowledge and Co-Action

Collective action requires skilled cooperative action, and this means being able to understand the communicative situation and know how and when to respond appropriately for the purpose at hand. This skill is of the performance of knowledge in co-action (Gill and Borchers, 2003), situated in the formation of legitimate peripheral participation (Lave and Wenger, 1991), and is a form of social intelligence for sustainable interaction.

The conception of 'knowledge', in this discussion, is rooted in praxis. In the 1980s, the concept of human knowledge and skill was reducible to logic and empirical validity for the purposes of designing information systems that could either substitute for the expert or support the expert in their everyday work. The 'knowledge engineer' was set the task to elicit and transform such skill into codable entities and structures. The notion of human-computer interaction was rooted in this

[1] Discussion with Duska Rosenberg, Royal Holloway.

[2] SANE project (2003), applied the concept of overhearing to the organizational context.

picture, leading some, such as Weizenbaum (1984) and Winograd (1986), to question it's basis. With the advance of interaction affordances with new technologies, the scope for thinking about human skill has shifted from its former cognitive focus, to one of context, location, and situation. This is a shift from a picture of knowledge as representation, to knowledge as emerging in the performance of representation. The emphasis is subtly moving from 'usability' to 'interaction', thereby placing an onus on the technology to be useful and invisible for human purpose.

This movement of conceptualizing the affordances of technology within human practices, is significant for organizations, which are increasingly facing distributed and mobile structures. Building technological affordances to share the 'tacit' knowledge of work practices, taken for granted in collocated settings, are still in development. Fruchter's work on zones of engagement provides an analysis of how engagement is affected and shaped by the physical configuration, the communication, and the technology interface, in tasks involving apprenticeship in distributed and collocated teams. Fukuhara, Chikama, and Nishida's work on building community support systems for community networks emphasizes the need for community analyses and an understanding of maintenance of the network, in order for any support system to be effectively used in truly creating sustainable community networks. Notsu, Katai, Kawakami and Shiose's work on developing agent systems for 'caring interaction' shows how theories of psychological and health care can be brought into agent design for human need and purpose. These writings, and others in this collection, seek to understand human practices in shaping systems design that are socially sustainable. These works are concerned with the processes of grounding in communication (Clark) within co-located and also distributed settings, where 'distributed' can be taken to mean space-time gaps in collocated as well as non-collocated space. With an increasing dependency upon technology to share information, organizations need to understand how people engage both directly and indirectly, through active participation and 'peripheral participation' or awareness, ['overhearers' e.g. SANE project, 2003].

This ability to transition these states of participation skillfully is part of our everyday communication, and we are aware of when we do not transition them well. For example, imagine that a meeting is underway and you want to make a suggestion, but your attempts are ignored by the other members. This is about the legitimacy to actively participate. Legitimacy is arrived at through a process of grounding in communication. This involves an incremental building of familiarity and understanding of each other at multiple levels, and is fundamental to co-ordination and collaboration (Clark & Brennan, 1991). Within the context of social intelligence, our representations of action, such as a gesture, a word, or an artifact, mediates knowledge of our feelings, ideas, intentions, states, beliefs, values. Learning to grasp these is a key element of the processes of grounding that take place in any collaborative work setting. Understanding how technology may best be used to facilitate this grounding, is central to this chapter. In order to do so, understanding how knowledge is grounded in communicative practices within organizations is essential.

A further dimension to 'grounding' in organizational knowledge is the role of imagination. This enables people to align their actions over space and time bringing together the past, present, and projected future, and is touched upon in the examples presented here. For example, if one person in a team leaves the organization, another member of that team may be able to take over and continue the path of the work of that member who has left, such that the identity of the knowledge that is produced remains coherent. This is common in organizations, and is marked in small organizations or firms, such as in architecture. This alignement in knowledge production and reproduction depends on this dimension of grounding. The theoretical link between this form of 'grounding' over space and time, and the immediacy of interactional grounding covered above needs to be developed. It is proposed that the essential elements are the same but that their relationship will only be salient when both situations are better understood as to how one informs the fuller nature of the other[3].

[3]This is work in progress and is being developed in a forthcoming paper by Gill.

3 Grounding and the tacit dimension

Grounding is enabled by the the tacit and explicit dimensions of knowledge. The tacit dimension of organisational knowledge has long been of concern in knowledge management research, as it is the kernal of knowledge formation and innovation, as well as constituting organizational identity and practices (Von Grogh, Ichijo, and Nonaka, 2000). As organizations become distributed in geographical time and space, grasping this tacit dimension by understanding its operation with the explicit is important if we are to develop technologies that can harness this transition in grounding effectively.

We acquire the tacit dimension 'through our socio-cultural experiences and working practices' (Gill, 1988). In On certainty (1969), Wittgenstein said that 'practice has to speak for itself' because 'rules leave loopholes open'. In other words, the utterance or gesture is not sufficient to explain the meaning behind itself. Only experienced persons can apply rules correctly and effectively; 'only practice gives words their meaning' (On Certainty, p.317). Johannessen (1986) further clarifies this notion of practice when he points out, 'the mastery of the practice is the tacit expression of a kind of understanding (vor-wissen) shown in correctly understanding the sayings and doings that make up the practice.' Practice, doing and acting lead to an understanding of concepts and rules (Gill op.cit. 325). It is in this context, that grasping the representations of action (e.g. words and gestures) is learnt in practice, and through grounding this we shape our understanding of the concepts and conventions we are expressing. 'It is necessary to have an experience of relationships... of fitting things, e.g. words, into context(s), which in turn lead to a 'familiarity' and, eventually, allow one to get a picture of their life'. (Gill opt.cit. 325). It is in this concept of the tacit and in experiencing, that imagination is both a part of the direct interaction as well as of the peripheral.

4 Skilled cooperative action

In this chapter we focus our analysis on how we engage with the representations of another's actions and move with these representations when interacting in a joint activity. Particular emphasis is on the rhythmic patterns in this coordination, and observations of how the use of artifacts can influence these. The motivation behind this focus is to understand the formation and transformation of knowledge in communication. Skilled cooperative action seems to depend on specific types of behavioural alignments between actors in an environment. These alignments allow for a degree of coordination or resonance between actors that constitutes the knowledge they have available to them, both explicitly and implicitly. Knowledge is seen here as a process that is dynamically represented in actors' behaviours and the tools, technologies, and other artifacts within an environment. These behaviours involve the sense of touch, sound, smell and vision. Coordinated structures of these behaviours and artifacts, structures which can be extended and transformed through technology, form the interactional space in which the process of knowledge creation through co-action takes place.

This analysis of coordinated structures in skilled cooperative action involves understanding what constitutes socially intelligent behaviour that we take so much for granted in our everyday behaviour. As a framework, the idea of Social Intelligence has been defined[4] (Nishida, 2003) as the ability of a collection of actors/agents to learn and to solve problems as a function of social structure, and to manage their relationships with each other through intelligent actions. Our study of such intelligent action is presented in this paper through a study of Body Moves (kinesics and kinaesthetics) within joint activities that involve negotiation and collaboration. Previous work on the metacommunicative[5] body dynamics of the engagement space (Gill, Kawamori, Katagiri, Shimojima, 2000) showed that composite dialogue acts, of gesture, speech and silence (Body Moves

[4]Toyoaki Nishida's summary of the idea of Social Intelligence at the close of the Social Intelligence Design Conference, July 8^{th}, Royal Holloway, London, 2003.

[5]Scheflen (1975) on the relation between kinesics (movements) and language: the former can 'qualify or give instructions' about the latter in a relation that Bateson called metacommunicative, whereby the 'movement of the body helps in clarifying meaning by supplementing features of the structure of language'. (p.11). Body Moves do just this, and contribute further to the idea that the structure of language lies in its performance.

(BM)), play an important role in the flow of information in interaction. The movements take place within shifting spaces of engagement. An *engagement space* is the aggregate of the participants' body fields of engagement, where an engagement field is based on some commitment in being bodily together. Within this field, there is a spatial fluidity of zones (reflection, negotiation, and action) of interaction (Gill, 2003)[6].

In further work on tacit knowing and the body, Gill (2004) shows how patterns of behavioural alignements change as the interaction moves from the state of information flow to knowledge transformation. This is seen as a process of grounding (Clark, 1996), where grounding occurs within the integration of the axes of awareness, activity and cognitivity of the tacit and explicit dimensions of human knowing (Gill, J. 2001). Within this process, the coordination of autonomous actions of individuals[7], termed Parallel Coordinated Action, plays a salient function. Managing the flow and formation of knowledge in co-action (Gill, 2003) is a form of social intelligence, and makes for sustainable communication, creativity, and innovation. It is culturally shaped.

The detailed understanding of interpersonal actions exhibited during the course of computer-mediated or conventional communication environments is part of the search for better design of communication media in organisations.

5 The Studies

The conceptual framework of knowledge in co-action is developed through a discussion about cooperative behaviour within collaborative and highly interactive design tasks. These involve participants producing conceptual designs together. We compare how acting upon various interfaces that have differing affordances, influences the communicative strategies they use to collaborate with each other. The qualitative analyses are of an experiment and an observation study of collaborative activity taking place in pairs and groups. We have sought to capture the management of both the body and speech spaces in the performance of knowledge in co-action.

The experiment[8] involved two drawing surfaces, used by different sets of subjects, a whiteboard and a SMARTBoard. The latter is a large-scale computer-based graphical user interface, and is touch-sensitive (an electronic whiteboard). 'Smart' technology does not permit two people to touch the screen at the same time. The contrast between drawing at the 'smart' and the whiteboard was expected to reveal whether or not there are particular differences in body moves and gesture and speech coordination at these interfaces. Preliminary analyses[9] revealed that the participants' commitment, politeness, and attention to each other was reduced at a single SMARTboard, showing behaviours that are in marked contrast to those of users at a whiteboard. Furthermore, the quality of the resulting design is lower when using the SmartBoard[10]. The focus of the analysis is of the gestures participants use to manage their interactions with each other and the interface during collaborative activity. This activity takes place in the 'iRoom', which is the laboratory of the Stanford Interactive Workspaces project[11] (Guimbretiere, Stone, and Winograd, 2001).

[6] The idenitification of the zones first emerged when reflecting out loud with Renate Fruchter, Stanford, in Spring 2002, about Donald Schon's work, in relation to this study.

[7] In his work on Knowledge Creation as the Production of Individual Autonomy, Saari (1999) discusses how in the world of mass media communications, the individual's capacity of subjective interpretation and tacit knowledge can drive creativity, and if recognised, can lead to the design of adaptable systems for the user. In work being developed above, Gill is finding that the management of autonomy is spontaneously conducted by individuals. In simultaneous autonomous action, she proposes they are able to arrive at tacit understanding of each other's positions in the communicative situation (Gill, 2004).

[8] Gill, S.P., Sethi, R., Martin, S. (2001), The Engagement Space and Gestural Coordination. In Proceedings, Gestualite et Oralite, Aix-en-Provence, France, 2001.

[9] For a preliminary discussion of this research, see: Borchers, J.O., Gill, S.P., and To, T. (2002). Multiple Large-Scale Displays for Collocated Team Work: Study and Recommendations, Technical Report, Stanford University.

[10] This could in part be due to the awkwardness of the interface for producing smooth drawings.

[11] `http://graphics.stanford.edu/projects/iwork/`

6 Parallel Coordinated Moves and collaborative activity

Whilst exploring the affordances of SMARTboard technology, we found that their inability to allow for parallel actions at the surface of the task being undertaken by participants, made for a useful case bed of data to analyse how such action affords collaborative activity. In the process of our observations, we also discovered that the use of multiple SMARTboards with groups of users produced similar patterns of behavioural alignments of Body Moves and Parallel actions[12], but in a very different form, to those undertaken at a whiteboard or over a sheet of paper on a horizontal surface (table). In the case of the SMARTboard, when one has to wait one's turn to act at the surface, it may (a) take longer to build up the experiential knowledge of that surface than if one could move onto it when one needs to, and (b) there is a time lag for the other person working with you to experience with you, in a manner of speaking, your experience of the situation, i.e. there is an awareness lag. The former and the latter difficulties are, we suggest, linked because of this experiential dimension of tacit or implicit knowledge. With multiple SMARTboards in parallel use, awareness of the experience seemed more fluid than that of a dyad at the one board, evident in the movements around the boards to gather and disperse where rhythms in behavioural alignment were halted. In the case of the single SMARTboard, the recourse is to use body field disturbance strategies to intercept and take the turn or force it.

The SmartBoard makes those actions that are invisible, or are extensions of ourselves, when acting at a whiteboard, visible. Winograd and Flores (1986) speak of how 'structure' in communication 'becomes visible only when there is some kind of breakdown' (Ibid.:p.68). For instance, acts of rubbing something out whilst another is drawing, checking something by pointing on it, or touching the surface with a finger or hand to think about an idea, etc. When all these aspects are inhibited or have to be negotiated, the fluidity of sharing an engagement space in an interactive drawing task becomes altered by the kinds of communication strategies available to participants to achieve collaborative activity. The simultaneous synchrony of parallel coordinated movement in drawing or being able to touch the surface together provides for a certain kind of awareness of states of contact within an engagement space. This synchrony allows for the multi-dimensional expression of ideas using a pen, hand or finger, to sketch ideas. These ideas can be rubbed out, located in one's self and made clear for the other person. Contact with each other's ideas can be made with gestures as well as speech. This is tacit bodily knowledge of self and intra-self.

Acting in parallel, e.g. drawing on the surface at the same, involves a degree of autonomy that is coordinated and simultaneous, and how participants manage it is cultural. In our study we observe patterns of movement from sequential (e.g. turn taking) to parallel actions and suggest that coordinated autonomous action is part of sustainable collaborative activity and knowledge creation[13]. The disruptive behaviours resulting from the inhibition of this simultaneous action, reveals the strain of not being able to use the cues of peripheral legitimacy that allow for it.

We have noticed that in group activity using multiple SmartBoards, problems being faced by any member(s) of the group are noticed, either by pauses in body action, or by someone saying something. The other group members then migrate towards the problem space and try to help resolve it. As they cluster around a board, they frequently take-turns by moving in and out of the problem focal spot to the outer rim, and once the problem is solved, they disperse to the separate boards. This interactional dance happens when all the participants are at the boards, as this enables an awareness that allows them the fluidity of movement within each other's problem spaces. The fact that there are four students to three boards, may help this fluidity as one person will, at any time, not be acting on the surface, and hence have a sense of actions occurring elsewhere. The proximity of participants to each other within the space facilitates overhearing[14] and the formation of legitimate peripheral participation which is significant in helping the awareness and maintainance of the engagement space of the group.

[12] See Borchers, J., Gill, S.P., To, T. (2002). Op cit.

[13] The relationship between autonomy and sustainable collaboration is being developed in a forthcoming paper by the author.

[14] People who overhear others talking in work environments are participants of a kind within that space and constitute part of the organisational knowledge (SANE project, Royal Holloway, London, 2003).

7 Conclusion

Collaboration and cooperation in joint activities is analysed as having three basic elements: the skill to grasp and respond to the representations of the tacit dimension of our actions (e.g. in body moves, gestures, sounds); the ability to coordinate this grasping and responding in a rhythmic synchrony of sequential and parallel actions; and simultaneous and coordinated autonomy that occurs within parallel coordinated movements, and involves awareness and attendance to the state of engagement in the space between us and interfaces.

This study of the complexity of body moves of pairs and groups of designers collaborating and cooperating on design tasks shows how they learn and solve problems as a function of communicative and social structure, and manage their relationships with each other through intelligent actions. By intelligence, we mean the skill of grasping and responding to the representations of the tacit dimension of our actions and knowledge, appropriately for the purpose at hand. We call this the performance of knowledge in co-action. A challenge for designing mediating interfaces is for them to afford us our human skills of engaging with each other, communicating information, and forming knowledge.

References

Allwood, J., Nivre, J., & Ahlsen, E. (1991). On the Semantics and Pragmatics of Linguistic Feedback. Gothenburg Papers. *Theoretical Linguistics,* 64.

Bateson, G. (1955). "The Message. 'This is the Play"'.In B. Schaffner (ed.*),Group Processes. Vol.II.* New York: Macy.

Bavelas, J.B. (1994). Gestures as part of speech: methodological implications. *Research on Language and Social Interaction.* 27(3), 201-221.

Birdwhistle, R.L. (1970). Kinesics and Context. University of Pennsylvania.

Borchers, J., Gill, S., and To, T (2002). Multiple Large-Scale Displays for Collocated Team Work: Study and Recommendations. *Technical Report. Stanford University.*

Clark, H.H. and Schaefer, E.F (1989). Contributing to discourse. *Cognitive Science,* 13, 259-294.

Clark, H.H. (1996). *Using Language.* Cambridge: Cambridge University Press.

Engel, R. (1998). Not Channels but composite signals: speech, gesture, diagrams and object demonstrations are integrated in multi-modal explanations. In M.A. Gernsbacher & S.J. Derry (Eds.) *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

Gill, S.P. (2004, in print). Body Moves and Tacit Knowing. To appear in Gorayska, B. and Mey, J.L. (Eds.) *Cognition and Technology: Co-existence, Convergence, Co-evolution.* John Benjamin. Publication due in August 2004.

Gill, S.P. and Borchers, J. (2003). Knowledge in Co-Action: Social Intelligence in Collaborative Design Activity. *AI & Society*, Vol.17.3.

Gill, S.P. (2002) *The Parallel Coordinated Move: Case of a Conceptual Drawing Task.* CKIR, Helsinki.

Gill, S.P., Kawamori, M., Katagiri, Y., Shimojima, A. (2000).The Role of Body Moves in Dialogue. *RASK*, Vol.12, pp.89-114.

Goodwin, C. (2003). Pointing as Situated Practice. In *Pointing: Where Language, Culture and Cognition Meet*, (ed.) Sotaro Kita, Lawrence Erlbaum.

Guimbretiere, F., Stone, M., Winograd, T. (2001). Stick it on the Wall: A Metaphor for Interaction with Large Displays. Submitted to *Computer Graphics (SIGGRAPH 2001 Proceedings).*

Kendon, A.(1970). Movement Coordination in Social Interaction:Some examples described. *Acta Psychologia*, 32.

Scheflen, A.E. (1975). *How Behaviour Means*. New York: Anchor Books.

Polanyi, M. (1964). Personal Knowledge: Towards a post critical philosophy. N.Y.: Harper and Row.

Saari, T. (1999). *Knowledge Creation as the Production of Individual Autonomy: How news influences subjective reality*. University of Tampere, Finland.

Winograd, T. and Flores, F. (1986). *Understanding Computers and Cognition. A New Foundation for Design*. Norwood, NJ: Ablex Corporation.

Close Encounters of the Virtual Kind:* Exploring Place-based Presence

G. Henri ter Hofte, Ingrid Mulder, Carla Verwijs
Telematica Instituut
P.O. Box 589, 7500 AN Enschede, The Netherlands
{Henri.terHofte|Ingrid.Mulder|Carla.Verwijs}@telin.nl

Abstract

Use of Presence and Instant Messaging (PIM) applications has grown very rapidly recently, not only at home, but also at work. Early studies on the use of PIM applications in the workplace, however, indicate that PIM applications need to be adapted towards the workplace context. In our research, we explore such adaptations, towards place-based presence systems, i.e., presence systems that are not only able to answer people-oriented presence queries such as "Who is online?" next to place-oriented presence queries, such as "Who is near?".
In this paper, we describe a design space for place-based presence systems, in which we identify the most important aspects and options that designers of place-based presence systems need to consider. We also report on our exploratory research in this design space, comprising two cycles of designing, implementing and evaluating place-based presence prototypes. We conclude with lessons learned for future research in place-based presence systems.

Keywords: Presence, awareness, Instant Messaging, virtual location, CoCoBrowse

1 Introduction

Presence technology is conquering the Internet in a rapid pace. The first generation of this technology is used in PIM applications such as AOL Instant Messenger (AIM), MSN Messenger and Yahoo! Messenger and answers questions such as "Who is online?", "Is person X online?" and "What is person X doing?". For millions of users [1], keeping in touch via PIM with friends and family has become part of their daily life. In March 2001, 74% of online teenagers in the US used PIM applications, versus 44% of online adults, according to a recent study (Lenhart et al. 2001).

Using presence information, one can observe that a particular person is likely to be available for communication, and the PIM application offers a lightweight means to initiate an IM conversation. This person, in return for giving up some privacy, hopes to be contacted only at suitable moments, can screen incoming messages, can plausibly deny presence by not responding (Nardi et al., 2000) and can respond later, simply by typing into the conversation window at a more suitable time. Continuous information about the presence of others and largely automatic derivation of presence information (as opposed to manual updates) make that PIM applications are more like a collaborative virtual environment (CVE) that *augments* physical reality rather than an immersive 3D CVE that *replaces* physical reality.

*Thanks go to William van Dieten, Henk Eertink Marjan Grootveld, Hans C.J. Kruse, Robert Slagter, Martin Snijders, and other members of the GigaCSCW project, to the Distributed Systems group at Ulm University and the members of Inter-*Actief* for their contributions. This work is sponsored by the Dutch Ministry of Economic Affairs as part of GigaPort (www.gigaport.nl), by SURFnet and by Cisco Systems.

[1] Recent figures (Jupiter Research, 2002) show that over 72 million unique users (persons, not accounts) at home and work in the US used PIM at least once during April 2002. That amounts to over 25% of the entire US population.

PIM applications are beginning to move into the workplace. In the US, the total time all users at work spent actively using PIM applications grew 110% between 9/2000 and 9/2001; the average PIM user at work spent 6.1 hours in September 2001 (Jupiter Research, 2001). Nevertheless, there were still about four times as many PIM users at home compared to PIM users at work. We believe that the nature of relations with one's co-workers differs significantly from the nature of relations with one's friends and family, which warrants investigation whether and how the design of PIM applications should be adapted towards workplace use. In our research, we explore presence mechanisms that allow exchange of presence information with a certain subset of people (e.g., a particular project team) not *always* (as in current PIM applications), but only *sometimes,* depending on real-time context information that can be derived from virtual places people visit during their work, which gives rise to a new model for presence we coin *place-based presence.*

In this paper, we report findings from an exploratory study into the use of place-based presence information in workplace PIM applications. First, we present an overview of the most salient features of current presence functionality in PIM applications. Next, we motivate the need for more extended presence information in the workplace. Then, we describe a design space for place-based presence systems, in which we identify the most important aspects and options that designers of place-based presence systems need to consider. We also report on our exploratory research in this design space, comprising two cycles of designing, implementing and evaluating place-based presence prototypes. We conclude with lessons learned for future research in place-based presence systems.

2 Presence, awareness and Instant Messaging

2.1 Instant Messaging

Like chat, instant messaging (IM) provides computer-mediated text-based near-synchronous communication (see Figure 1). After a sender types a message and hits the enter key or clicks a "Send" button, the message, preceded with the display name of the sender and possibly a timestamp, gets appended as the latest entry in the conversation window, not only on the sender's screen, but also on the receiver's screen, usually within fractions of a second. IM has both characteristics of e-mail and of telephony. Like e-mail messages, instant messages are readable and reviewable, which affords self-paced reading and having multiple conversations at the same time, something most users find very hard to do with telephony. Unlike in e-mail, where the thread of conversation is usually constructed and reconstructed with quotes from the original message interspersed with replies, the context of an instant message is typically found only in earlier contributions in the conversation. IM is immediate like telephony, but it lacks non-verbal cues like intonation, which can be compensated to some extent with emoticons such as :-) and . Some IM systems provide features that regulate turn-taking, such as showing when a sender starts typing, or by showing letters as soon as they are typed by a sender (Tang et al., 2001). Some systems support persistence: parties joining can review parts of the conversation that occurred before they joined.

2.2 Getting into a Conversation

A distinctive feature of instant messaging compared to chat is the way people get into a conversation and the role that presence functionality plays. In chat systems, parties *join* a pre-existing chat room (also known as a "channel" or "topic"). Once inside, and only when inside a chat room, a party can become aware of the presence of other parties and get notifications of other parties joining and leaving the room. With chat, joining a room implies being available for conversation. In instant messaging systems, a user can *initiate* a conversation with another user (with a first message) and in some systems a user can *invite other parties* to join a conversation. Upon receiving a first instant message or an invitation to join a conversation, conversational availability has not yet been negotiated with the receiver. Hence, like a telephone call, an instant message can be very interruptive.

Figure 1: A typical Instant Messaging window

2.3 Presence and awareness

In PIM applications, presence information of a receiver is shown to the sender continuously and changes are notified, thus making the sender aware, not only during a conversation (as in chat systems), but also *before* a conversation is initiated (see Figure 2). Though this may seem a subtle distinction, this very feature makes PIM applications "socially translucent" (Erickson and Kellogg, 2000), which seems to be crucial to negotiating conversational availability (Nardi et al., 2000).

To illustrate the concept of social translucence, Erickson and Kellogg (2000) present the problem of an opaque door in a hallway that sometimes is smashed inadvertently into another person's face when it is opened just at the moment when that person is approaching the door. One design to solve this problem is to post a note on the door telling people to open the door with caution. Another design (which adheres to the principle of social translucence) is to use a glass (window in the) door: now the person opening the door (person A) can see the other person (B) approaching, which helps to reduce the problem, not only because A is *aware* of B approaching, but also because social norms typically make A feel *accountable*: A is aware that B is aware that A is aware that B is approaching.

Just like the glass window in the door does not require additional effort of people approaching the door to make others aware of their presence, most current presence systems can automatically derive some presence information, such as "online", "offline", and "away", from user activity. Turning on the computer and connecting to the internet sets the status to online; a configurable duration of keyboard and mouse inactivity implies being away; new keyboard or mouse activity implies being online; disconnecting from the internet or turning off the computer implies being offline. Other presence information, such as busy, on-the-phone, out-to-lunch, be-right-back, requires user effort. Some systems allow a user that is logged on to appear offline.

So, current presence functionality in PIM applications can answer the questions "Who is online?", "Is person X online?" and to a lesser extent: "What is person X doing?". Using this presence information, a sender can observe that a particular person is likely to be available for communication, and the PIM application offers a lightweight means to initiate an IM conversation. The receiver, in exchange for giving up some privacy, hopes to be contacted only at suitable moments, can screen incoming messages, can plausibly deny being present by not responding (Nardi et al., 2000) and can respond later, simply by typing into the conversation window at a more

Figure 2: A typical window showing presence

suitable time.

2.4 Establishing Trust with "Buddies"

Instead of allowing everybody to see each other's presence information at all times, which would run into serious privacy, information overload and technical problems, most presence systems currently apply a much more restrictive trust model. User A explicitly has to request user B permission to subscribe to his presence information. This process is usually reciprocal: after user B has granted user A permission, they are "buddies": not only is user B added to the "buddy list" of user A, but also is user A added to the buddy list of user B. The presence information of all buddies is shown in the buddy list window. Initiating instant messaging conversations with other users is only possible for users that are on your buddy list. In some presence systems a person can be "blocked": his rights are temporarily and unilaterally revoked.

3 Towards Workplace Presence

Despite preliminary evidence for the utility of PIM at the workplace (Isaacs et al., 2002; Jupiter Research, 2001), the introduction of PIM applications in the workplace is not guaranteed to be successful. An early adoption study of a chat application in the workplace (Bradner et al. (1999)) found its use to be "healthy" in some subgroups but "fragile" in other subgroups. In a more recent study, Herbsleb et al. (2002) found that contrary to their expectations, introduction of PIM applications in the workplace proved difficult and adoption only slightly improved after modifications to the tool, to default permissions and to the way the tool was introduced in the organization. These results warrant further investigation whether and how the design of PIM applications should be adapted towards workplace use.

As we describe below, the nature of relations and interactions with one's co-workers differs significantly from the nature of relations and interactions with one's friends and family.

A first observation is that the way of establishing trust scales poorly with respect to the number of users that want to establish buddy relations: to establish full trust between 4 persons, 4*3/2=6 bilateral agreements need to be established. Between 10 persons, already $10\times9/2=45$ bilateral

agreements need to be established. This motivated some designers of workplace PIM systems (e.g., (Herbsleb et al., 2002)) to choose a group-based trust model: if you join a trust group, then all members of the trust group are on your buddy list and you are on theirs, which scales much better.

A second observation is that the trust model is very crude: either you establish a trust relation with someone (or with a group), in which case that person (/ the other members of that group) and you can *always* observe each other's presence information, or you don't establish a trust relation, in which case that person (/ the other members of that group) and you can *never* see each other's presence information and cannot engage in instant messaging conversations. This might not be very problematic when dealing with a small set of family and friends and with whom you expect to resolve unwanted interruptions easily. The trust model seems very crude when dealing with e.g., a larger set of co-workers in a multi-project environment. Of course, one could use multiple identities, or establish and tear down or block and unblock trust relations depending on, e.g., the project one is currently working on, but we believe this is too cumbersome for most workplace users. Another trust model emerges from community websites. Many community-support systems, "even inexpensive discussion boards, now have a list of who is on" (Wenger, 2001, p.47). The presence model of these community-support systems consists of a group-based trust model; presence information not only indicates "who is online" but also "who is here", i.e., who is currently logged on to the community website. This presence model may work well for communities that use a community website as their primary resource. However, many groups and communities in workplaces often use a variety of such resources, including shared network drives, intranet/extranet/internet websites, newsgroups, etc.

A third and final observation is that in workplace environments, answering the question "Who is online" may not provide much added value, e.g., when co-workers are almost always online at the same time. In such cases – as well as in the future when more and more people use always-on Internet access technologies, such as cable, ADSL, and GPRS – more detailed presence information might be needed than just online/offline status.

4 Place-based Presence: A Design Space

Place-based presence systems can not only provide answers to questions about people, but also about the following type of questions about *places* (such as pages on a website)*:* "Who is here/near?", "Where is person X?", and "What is person X doing there?" These questions can be answered based on a combination of the trust relation between two users, their presence locations, activities at these locations and presence and awareness scopes.

The idea to visualize presence of people on the World Wide Web and to use that as a basis for chance encounters and real-time communication can at least be traced back to systems such as WebTalk from the Sociable Web project (Donath and Robertson, 1994) and the Virtual Places platform from Ubique (Shapiro, 1994), a company later acquired by the Lotus (which again is now part of IBM), which now offers place-based awareness in their IBM Lotus Instant Messaging product (Lotus Development Corporation, 2001). Some PIM applications, such as Odigo[2] (Odigo, 2002) can show which other Odigo users are on the same page or the same website. In these tools, presence locations are not laid out in a space and there is no virtual distance between places. CoBrow (Sidler et al., 1997) is one of the first tools that support virtual distance between web pages as presence locations, based on the number of hyperlinks that must be traversed between web pages. To our knowledge, our place-based presence tools, described in the section "Exploring Place-based Presence", are the first place-based presence tools to support activities on WebDAV-based websites and/or document servers (Whitehead and Goland, 1999) as presence information ("locking here").

Below, we describe a design space for place-based presence systems, i.e., a framework that identifies various aspects designers need to consider when designing a place-based presence system. We will relate various concepts in the design space to concepts from the spatial model of interaction (Benford and Fahlén, 1993).

[2]On May, 2002 Odigo was aquired by Comverse

4.1 Trust Model

Some presence systems allow anyone who has access to a presence server to see presence information of others (e.g., CoBrow), other systems are more restrictive (like most PIM applications that only allow buddies to see presence information). The establishment of trust can be compared to the collisions of auras in the spatial model of interaction (Benford and Fahlén (1993)). We distinguish four aspects:

- Opt-in / opt-out / fixed: In an opt-in trust model, others cannot see your presence information, unless you explicitly give them permission. In an opt-out model, others can see your presence information, unless you explicitly denied them permission. A special case is a fixed model, where an administrator instead of the end users determines who can see presence information.

- Bilateral / group: In a bilateral trust model, you give or deny each person rights to your presence information separately. In a group trust model, you give or deny a group rights to see presence information.

- Reciprocal / non-reciprocal: In a reciprocal trust model, if person A (/group G) has the right to your presence information, you also have the rights to the presence information of A (/each member of G). In a non-reciprocal trust model, this may not apply.

- Permanent / blockable / place-based: In a permanent trust model, people can see your presence information as long as you gave them the rights to do so based on bilateral or group-wise arrangements. When you can temporarily block persons from seeing your presence information, we call the trust model blockable. If the right to see your presence information can be based on the presence location (see below), we call the trust model place-based. Place-based trust models allow you to e.g., only give project co-workers rights to see your presence information as long as you are browsing in a project website, editing a project document.

4.2 Presence Location and Virtual Distance

When users access shared resources, e.g., browse the web, edit files from a shared network drive, or read or post in newsgroups, in a sense, they are present at a location in cyberspace. "In a sense", because many assumptions we have when someone is present in physical space, such as being aware of someone's presence, and being able to initiate contact and communicate with that person, do not necessarily hold in cyberspace; many systems that provide access to shared resources are not socially translucent (Erickson and Kellogg, 2000).

In place-based presence systems, presence *location* information constitutes a primary form of presence information: not only the fact that a person is online (i.e., somewhere in cyberspace, without knowing where), but also which shared resource a person is accessing, constitutes presence information (e.g., *where* that person is in cyberspace) can be made available to trusted other persons. By relaying presence location information, systems can be made more socially translucent.

Presence location information is expressed by coordinates, e.g., URLs. Unlike in physical reality, users can be at multiple coordinates simultaneously (e.g., using multiple web browser windows). Presence location information is typically derived automatically when people access shared resources, thus providing for social translucence.

Place-based presence systems that only need to answer the questions "Who is here?" and "Where is person X?" typically only need to do equality tests or pass coordinates.

For place-based presence systems that also need to answer the question "Who is near?", it should be possible to calculate *virtual distance* between the coordinates. As explained in subsequent subsections, virtual distance can be used to determine who can and who cannot be seen in a list of currently present users. To calculate virtual distance, presence location coordinates need to be laid out in a space. Inspired by work from Dix et al. (2000), and CoBrow (Sidler et al., 1997), we distinguish three types of spaces and coordinates:

- Topological space, with coordinates such as <Netherlands. Enschede. University of Twente. KCT-building. H123> and <cscw. components. presence>. A topological space is organized as a hierarchy of spaces, that contain locations, which themselves can be spaces, etc. Distance in a physical topological space is expressed with terms such as "in the same room", "in an adjacent room", and "in the same building". Virtual distance is expressed with terms like "on the same web page", "on the same website", and "online".

- Graph space, which consists of nodes (e.g., intersections in a road network or pages in the WWW) and edges (e.g., roads between the intersections, hyperlinks). Virtual distance is expressed in terms like "4 blocks away", "2 clicks away".

- Cartesian space, with e.g., 3D coordinates such as <2,4,5>, for a space that is best characterized by of orthogonal dimensions. Virtual distance is expressed in terms like "within 1m."

One way to derive a presence location is to interpret the URLs of a website as a topological space of nested locations, e.g., interpreted as a coordinate in a topological space, `http://www.microsoft.com/net/whatis.asp` would be contained within the space `http://www.microsoft.com/net`, which again would be contained within the space `http://www.microsoft.com`. Another way to derive a presence location is to interpret a website as a graph space, with the pages as the nodes and the hyperlinks between the pages as the edges. However, not all URLs have a meaningful structure (e.g., URLs in the ACM Digital Library: `http://doi.acm.org/10.1145/344949.345004`) and the number of clicks may not always be the best way to measure a meaningful virtual distance.

We believe that a mechanism should be available for authors and administrators of shared resources (e.g., websites) that allows them to make their existing website better suitable as meeting place by laying out their website in a virtual presence space and associate each resource with a presence location within that space. One approach in this direction would be to use an automatic content-based semantic mapping that maps web pages that are semantically close to each other onto locations that are close to each other in one of the previously mentioned coordinate spaces.

4.3 Presence Scope(s)

A presence scope (similar to the "nimbus" in the spatial model of interaction Benford and Fahlén (1993)) specifies the maximum virtual distance at which another trusted user can observe particular presence information about a user (see Figure 3). One user may use multiple presence scopes, e.g., "people with me on the same website can see me, but cannot see where I am within the website" and "people on the same web page can see whether I am focusing on that page".

4.4 Awareness Scope(s)

An awareness scope (similar to the "focus" in the spatial model of interaction Benford and Fahlén (1993)) specifies the maximum virtual distance at which a user wants to get notifications about particular presence information of users that trust him (see Figure 3). One user may use multiple presence scopes, e.g., "I want to know where people are that are with me on the same web page" and "I want to see whether people with me on the same web page are focusing on the page".

4.5 Activity

What a user is doing at a location also constitutes presence information. For example, in addition to browsing a web page, this may also involve information whether the user is actually focusing on this page or not (since the user may have multiple windows open), whether the user is editing this page or not (which is relevant for a WebDAV-based website, where users can also edit web pages). We consider this a secondary type of presence information.

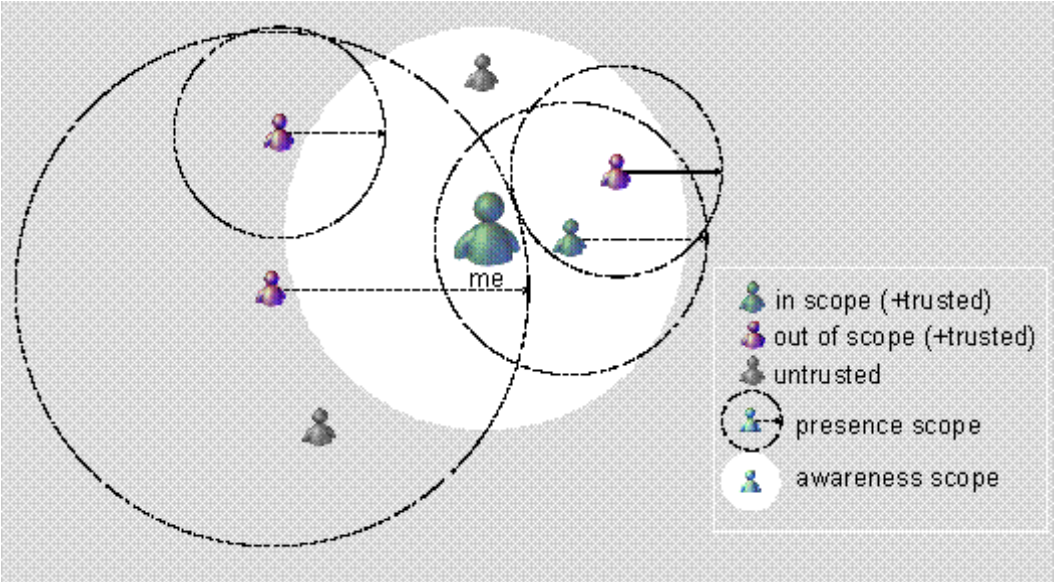


Figure 3: Conceptual representation of presence and awareness scopes

4.6 Presentation

For the presentation of place-based presence information, designers can choose basically between a 1.5D, 2D, and 3D presentation. A 1.5D presentation (list of users, grouped in a hierarchy, e.g., users on same page/site/anywhere) can be very concise. Representations in 2D and 3D can show both virtual distance and direction of the presence location of other users.

4.7 Presence Awareness Service

The questions "Who is near?", "Where is person X?", and "What is person X doing there?" can now be answered based on a combination of the trust relation between two users, their presence locations and their presence and awareness scopes. That is, a user B only appears in the presence list of a user A when the following applies:

- B trusts A to see his presence information;
- A's presence location is within B's presence scope;
- B's presence location is within A's awareness scope.

4.8 Using Concepts from the Design Space

The presence awareness service found in most current PIM applications can be characterized with the concepts from the place-based presence design space as follows:

- *Trust model*: opt-in, reciprocal, bilateral, blockable;
- *Presence location*: (does not apply);
- *Presence scope*: infinite
- *Awareness scope*: infinite
- *Activities*: online, offline, away (automatic), busy, on-the-phone, out-to-lunch, be-right-back (user indicated).
- *Presentation*: 1.5D

5 Exploring Place-based Presence

In this section we describe our exploratory research into place-based presence. First, we describe the design and implementation of a prototype of a place-based presence tool and we elaborate findings of the evaluation of this first prototype in a small user study. Then, we describe the redesign for a second prototype of a place-based presence tool and describe findings from the evaluation of the second prototype in a student association.

5.1 First Pilot Study

Inspired by the CoBrow system (Sidler et al., 1997) and research into informal and opportunistic communication (Whittaker et al., 1994), we designed and implemented the first prototype of CoCoBrowse (Kruse et al., 2000), an add-on to Microsoft Internet Explorer (see Figure 4) that can answer "who is here"-style presence questions for web browsers. CoCoBrowse was designed as one of the first components of our component groupware platform CoCoWare .NET (Slagter et al., 2002). On WebDAV-based websites, users can lock and edit documents. CoCoBrowse can detect such locks and present this as presence information. The presence awareness service of this first prototype of CoCoBrowse can be characterized as follows:

- *Trust model*: none (open for anyone);
- *Presence location*: URL of page browsed to;
- *Presence scope*: this URL (toggle button in "Visible" position) / none (toggle button in "Invisible" position);
- *Awareness scope*: this URL (toggle button in "Visible" position) / none (toggle button in "Invisible" position);
- *Activities*: focusing, defocusing (on web browser window; automatically detected), locking, unlocking (on WebDAV files; automatically detected);
- *Presentation*: 1D.

Users could start a real-time conference with any of the other users on the same web page, by double clicking on their name. This initiated a NetMeeting conference with that user (which provides audio and video conferencing, application sharing, shared whiteboard and file transfer).

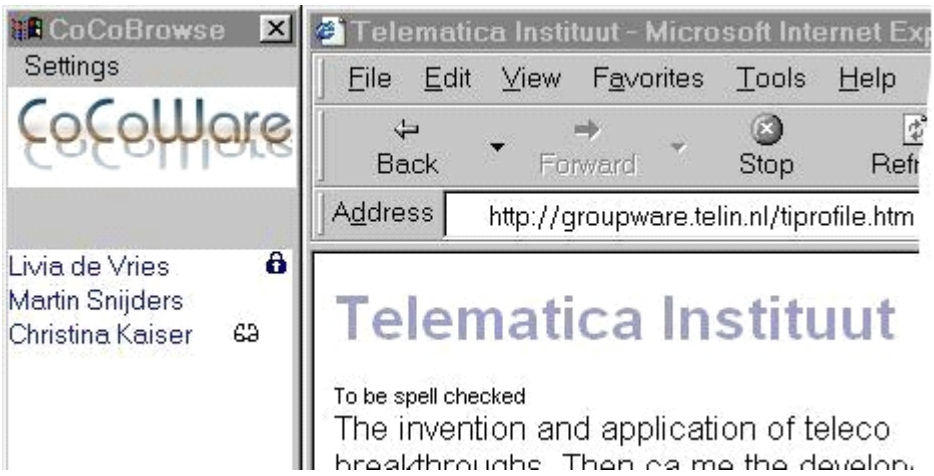


Figure 4: First CoCoBrowse prototype: 3 persons browse to the same file: Livia has a lock (does not focus), Martin reads (does not focus) and Christina reads (and does focus on) it.

In this first pilot study, we asked 17 people to install the first CoCoBrowse prototype, to use it in their daily project work during four weeks. All were member of the project team in which

CoCoBrowse was developed; a multidisciplinary team consisting of programmers, managers, social and technical researchers, and interaction designers. Only four were directly involved with the design and implementation of CoCoBrowse. Afterwards, we collected data with an online survey. Subsequently, we organized an evaluation meeting, in which we discussed the results of the online survey with all the participants. A small group of users decided to execute a stress test of the system by scheduling a six person meeting with CoCoBrowse at a pre-determined website. The users saved their chat-logs and made them available for our evaluation.

In the evaluation meeting, the respondents indicated they felt the critical mass needed to test this kind of tool was not reached. Furthermore, they suggested starting CoCoBrowse in combination with Internet Explorer. In the first prototype you had to start a special version of Internet Explorer, not the normal Internet Explorer. The suggested change would give users more trust in one's online status. After all, if a person's status was presented as "offline", this could also mean that the person did not start the special version of Internet Explorer with CoCoBrowse, although he was in fact browsing.

During the scheduled "stress testing" meeting, the respondents frequently lost track of each other and used other communication media (phone, other PIM applications). Based on the feedback from the survey and the evaluation meeting, we attribute this primarily to the "tunnel vision" that CoCoBrowse offered: you have to be at the exact same URL in order to see each other. People frequently resorted to asking "where are you"-type questions through other media. We also noticed that people assumed the system would allow people to browse together, e.g., by following another person, similar to the "navigate together" feature in many collaborative web browsers. Various improvements were suggested by users, e.g., widening the tunnel vision (e.g., "who is on this site"), in combination with showing where a particular person is; providing a "user X browsed to URL-Y" or a "navigate together" feature (by one user described as "stalking") and presence indications that only slowly fade when users left ("who was here recently?").

Also, during the evaluation meeting, differences were noted who could access presence information; CoCoBrowse was completely open, whereas most PIM applications access to presence information was regulated in an opt-in, bilateral fashion.

5.2 Second Pilot Study

Based on our evaluation, we changed the prototype (see Figure 5) and based it on the Virtual Presence System VPS of the University of Ulm (Christein, 2002). We added a page chat window that was always visible below the page being viewed and that could be used to chat with other persons also present at that URL. Unlike the chat in NetMeeting, which required several clicks to initiate, this chat was readily available. As a workaround for the problem of losing track of each other when a person browses away from a URL, we implemented a "My contacts" window that showed online/offline presence status of a fixed set of buddies – provided they had activated CoCoBrowse. We also added a *personal chat* feature (not shown in Figure 5), which could be initiated both from the "My contacts" and from the window that showed people present on the same URL.

The presence awareness service of the additional "My contacts" service can be characterized as follows:

- *Trust model*: fixed;
- *Presence location*: (does not apply);
- *Presence scope*: infinite;
- *Awareness scope*: infinite;
- *Activities*: online, offline (automatically detected when CoCoBrowse was activated/deactivated)
- *Presentation*: 1.5D

To circumvent the need to remember to start a special version of Internet Explorer to be able to use CoCoBrowse, the second CoCoBrowse prototype could be started from a button in the Internet Explorer toolbar.

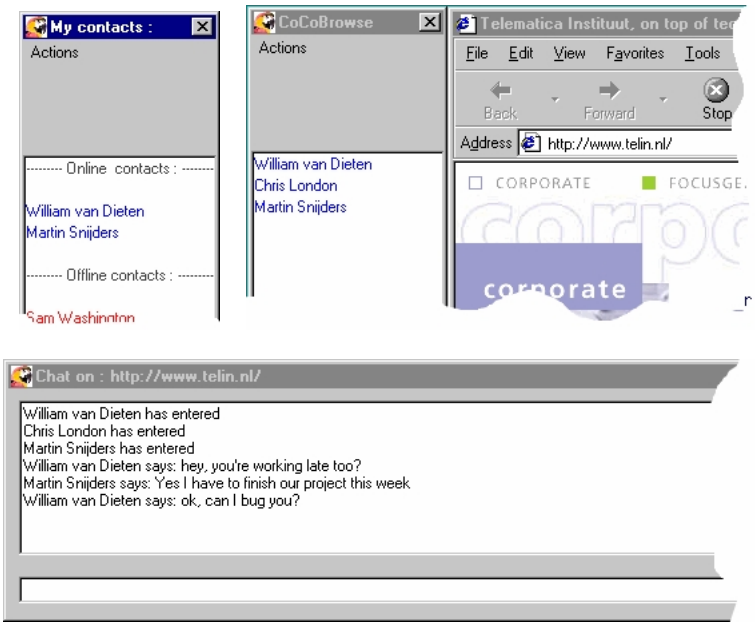


Figure 5: Second CoCoBrowse prototype: contacts William and Martin are online and on the same page, as well as non-contact Chris; contact Sam is offline.

In the second pilot study, we evaluated the second prototype of CoCoBrowse in a setting that did not involve people associated with the project in which the tool was made. We asked the members of various committees of Inter-*Actief*, a student association for computer-science students, to install and use CoCoBrowse for two weeks. We gave a short demonstration to the students on how the tool worked and which functionalities were available. After this introduction, the students received an e-mail with a link to download CoCoBrowse. In this way, we could see how many people downloaded the tool. We did not tell them how to use it, because this was part of the research: similar to a real situation people need to find out themselves what is the best way to use it and when to use it. Afterwards, we collected data with an online survey.

Due to their small number (N = 7), the students never met at a website by chance, but instead they made appointments for online meetings. Apparently, the critical mass of users required for chance encounters on the web was not reached. In retrospect, we feel that the fact that CoCoBrowse did not start the "My contacts" service (and announce the local user as being online) at system start up and that it did not start the "Who is here" window (and announce the local user as being here) by default when Internet Explorer was started, may also have hampered social translucence and trust in online status of others.

The respondents were experienced PIM users whose expectations were formed by earlier experience with PIM applications. It appears such people may not be prepared to spend time persuading their established contacts to use a new, incompatible system, if the benefits of that new system are unclear or too small. For the student association, instead of adding PIM features to a place-based presence system, it might have been better to add place-based presence features to existing PIM applications.

Table 1: PIM applications and our prototypes in terms of the design model

	Current PIM applications	Prototype Pilot 1	Prototype Pilot 2 (MyContacts)
Trust model	opt-in, reciprocal bilateral blockable	None	fixed
Presence location		URL	N/A
Presence scope	infinite	this URL/ none	infinite
Awareness scope	infinite	this URL/ none	infinite
Activities	online offline away busy on-the-phone out-to-lunch be-right-back	focusing, defocusing, locking, unlocking	online offline
Presentation	1.5 D	1D	1.5D

6 Conclusion

Recently, the use of presence and instant messaging applications has grown rapidly, not just at home, but also at work. However, early studies reveal that these PIM applications may have to be adapted to the workplace. We investigated whether place-based presence, i.e., presence enhanced with concepts from the spatial model of interaction, adds value to presence applications in workplace settings. We explored the design space for place-based presence applications with two designs and evaluations of place-based presence applications in real-life settings. Their features are summarized in Table 1.

From the evaluation of pilot studies with our place-based presence prototypes in our own project group and in an external student community, we learned the following lessons:

- Place-based presence applications should be designed as an extension of existing PIM applications, which allow people to exchange place-based presence information with some contacts, and stay backwards compatible with other contacts.

- (Place-based) presence systems should require very little user effort to update presence information. Not starting a presence system by default at system start up or a place-based presence system at start up of the place-visiting application (e.g., a web browser) already violates social translucence and may harm fidelity of presence information, which lowers trust in presence status.

- If a place-based presence system only shows other users at exactly the same URL, people hardly meet by chance and easily lose track of each other. Facilities are needed to avoid such "tunnel vision", e.g., with wider presence and awareness scopes for people to see each other.

With respect to the added value of place-based presence in general, we conclude that more work is needed to find out under what circumstances place-based presence systems can live up to the promise of providing additional useful information to discover and appropriate timing for communication with others. We look forward to pursue this research armed with the insights from this paper.

References

Benford, S., & Fahlén, L. (1993). A spatial model of interaction for large virtual environments. In G. De Michelis, C. Simone and K. Schmidt, editors, *ECSCW '93: Proceedings of the third European conference on computer supported cooperative work*, pages 109-124, 1993.

Bradner, E., Kellogg, W. A., & Erickson, T. (1999). The adoption and use of 'BABBLE': A field study of chat in the workplace. In S. Bodker, M. Kyng, and K. Schmidt, editors, *ECSCW '99: Proceedings of the sixth European conference on computer supported cooperative work,* pages 139-158, 1999. Retrieved from `http://www.pliant.org/personal/Tom_Erickson/AdoptionOfBabble.html`

Christein, H. (2002). The virtual presence system. Retrieved from `http://legoland.informatik.uni-ulm.de/virtual_presence/index.html`

Dix, A., Rodden, T., Davies, N., Trevor, J., Friday, A., & Palfreyman, K. (2000). Exploiting space and location as a design framework for interactive mobile systems. ACM Transactions on Computer-Human Interaction, 7, pages 285-321. Retrieved from `http://doi.acm.org/10.1145/355324.355325`

Donath, J., & Robertson, N. (1994). The sociable web. In *Electronic Proceedings of the Second World Wide Web Conference '94: Mosaic and the Web.* Retrieved from `http://archive.ncsa.uiuc.edu/SDG/IT94/Proceedings/CSCW/donath/SociableWeb.html`

Erickson, T. & Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. ACM Transactions on Computer-Human Interaction, 7, pages 59-83. Retrieved from `http://doi.acm.org/10.1145/344949.345004`

Herbsleb, J. D., Atkins, D. L., Boyer, D. G., Handel, M., & Finholt, T. A. (2002). Introducing instant messaging and chat into the workplace. In L. Terveen and D. Wixon, editors, *Proceedings of the CHI 2002 conference on human factors in computing systems,* pages 171-178, 2002. Retrieved from `http://doi.acm.org/10.1145/503376.503408`

Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D.J., & Kamm, C. (2002). The character, functions and styles of instant messaging in the workplace. In C. Neuwirth and T. Rodden, editors, *CSCW2002: Proceedings of the ACM 2002 Conference on Computer Supported Cooperative Work,* pages 11-20, 2002. Retrieved from `http://doi.acm.org/10.1145/587078.587081`

Jupiter Research. (2001). Total time spent using instant messaging jumps 110 percent at work and 48 percent at home versus last year, reports Jupiter Media Metrix. Jupiter Media Metrix Press Releases. Retrieved from `http://www.jmm.com/xp/jmm/press/2001/pr_111401.xml`

Jupiter Research. (2002). New Instant Messaging Application Poses Threat To Major Services, According To Media Metrix Internet Ratings. Jupiter Media Metrix Press Releases. Retrieved from `http://www.jmm.com/xp/jmm/press/2002/pr_052902.xml`

Kruse, H. C. J., Slagter, R. J., & Ter Hofte, G. H. (2000). Collaborative component software: The CoCoWare framework and its application (Rep. No. TI/SS/2000/033). Enschede, The Netherlands: Telematica Instituut. Retrieved from `https://doc.telin.nl/dscgi/ds.py/ViewProps/File-9506`

Lenhart, A., Rainie, L., & Lewis, O. (2001). Teenage Life Online: The rise of the instant-message generation and the Internet's impact on friendships and family relationships. Washington, D.C., USA: Pew Internet & American Life Project. Retrieved from `http://www.pewinternet.org/reports/toc.asp?Report=36`

Lotus Development Corporation (2001). Sametime 2.5 Java toolkit tutorial. Lotus developer domain website. Retrieved from `http://doc.notes.net/uafiles.nsf/0568d30ef4fb8f5e85256ac30068777d/$FILE/st25javatktutorial.pdf`

Nardi, B. A., Whittaker, S., & Bradner, E. (2000). Interaction and outeraction: Instant messaging in action. In S. Whittaker and W. A. Kellogg, editors, *CSCW2000: ACM 2000 Conference on Computer Supported Cooperative Work*, pages 79-88, 2000. Retrieved from `http://doi.acm.org/10.1145/358916.358975`

Odigo, I. (2002). Odigo: World's greatest instant messenger. Odigo website [On-line]. Retrieved from `http://www.odigo.org/features/windows.html`

Shapiro, E. (1994). Virtual places™: A foundation for human interaction. In *Electronic Proceedings of the Second World Wide Web Conference '94: Mosaic and the Web*. Retrieved from `http://archive.ncsa.uiuc.edu/SDG/IT94/Proceedings/WebProd/shapiro/demo1.html`

Sidler, G., Scott, A., & Wolf, H. (1997). Collaborative browsing in the World Wide Web. In *Proceedings of the 8th Joint European Networking Conference*, pages 122-1-122-8. Retrieved from `http://www.terena.nl/conf/jenc8/papers/122.ps`

Slagter, R. J., ter Hofte, G. H., & Kruse, H. C. J. (2002). The CoCoWare .NET architecture 2.0 (Rep. No. TI/RS/2002/123). Enschede, The Netherlands: Telematica Instituut. Retrieved from `https://doc.telin.nl/dscgi/ds.py/ViewProps/File-26932`

Tang, J. C., Yankelovich, N., Begole, J., Van Kleek, M., Li, F., & Bhalodia, J. (2001). ConNexus to Awarenex: Extending awareness to mobile users. In J. Jacko and A. Sears, editors, *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 221-228. Retrieved from `http://doi.acm.org/10.1145/365024.365105`

Wenger, E. (2001). Supporting communities of practice: a survey of community-oriented technologies. Homepage of Etienne Wenger. Retrieved from `http://www.ewenger.com/tech`

Whitehead, E. J., & Goland, Y. Y. (1999). WebDAV: A network protocol for remote collaborative authoring on the web. In S. Bodker, M. Kyng and K. Schmidt, editors, *ECSCW '99: Proceedings of the sixth European conference on computer supported cooperative work*, pages 291-310, 1999. Retrieved from `http://www.ics.uci.edu/~ejw/papers/dav-ecscw-A4.pdf`

Whittaker, S., Frohlich, D., & Daly-Jones, O. (1994). Informal workplace communication: what is it like and how might we support it? In B. Andelson, S. Dumais and J. Olson, editors, *Proceedings of the CHI '94 conference on human factors in computing systems*, pages 131-137. New York: ACM Press.

Agent-based Content Management System

Hidekazu Kubota
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 Japan
kubota@ii.ist.i.kyoto-u.ac.jp

Jaewon Hur
Graduate School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
jwhur@kc.t.u-tokyo.ac.jp

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

Abstract

This paper describes the management of huge heterogeneous content using the agent-based content management system. The knowledge card that can wrap up heterogeneous content is proposed. The knowledge channel that is a model of content management system is also proposed and implemented by using conversational agent and dynamic program table. An empirical experiment conducted in three communities implies that the knowledge channel allows the management of huge heterogeneous content.

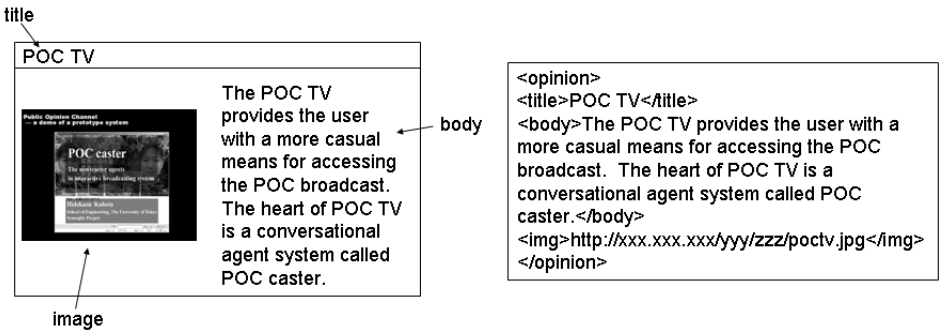
Keywords: personal content management system, conversational agent

1 Introduction

The purpose of this paper is a computational management for personal content creation. The personal content means here is a set of essays and papers that are created casually. The management of personal content is essential work for human intellectual life at present, for instance many people publish their personal journals on web pages. However, it is troublesome to manage huge personal content because the content can include heterogeneous parts (such as text, photos, music and movieclips), its constituent parts are different sizes, and moreover boundaries of its topics are not clear.

This paper is intended as the computational management of the unstructured heterogeneous content mentioned above. The essential idea is a knowledge card that is a well done piece of heterogeneous content. A knowledge card wraps up a few sentences of text and an image (or a movieclip) that represent one topic, so intra-card structure is too heterogeneous for a computer to understand, while inter-cards boundaries are very clear. Such distinction between cards is important clues for a computer to understand the whole structure of the content. In addition to that, these cards are easy to rearrange because their grain size is generally regular.

To manage knowledge card formed content, we have proposed a knowledge channel model. This is a model of an agent-based content management system that integrates the conversational process and the editorial process for managing content. These processes complementarily grow the content. The conversational process manages time series of content through conversation, while the editorial process manages spatial relations among content. To support the conversational process, we have proposed conversational agent that can talk with people about content, and



(a) The component of the knowledge card

(b) XML form of the knowledge card

Figure 1: The knowledge card

gather their opinions about it. On the other hand, we have proposed the landscape that enables people to grasp huge content spatially by using knowledge visualization.

2 Knowledge card approach

The knowledge card can include heterogeneous information, as a title sentence, a few body sentences and an image or a movieclip, nevertheless the card can be roughly retrieved by using text retrieving techniques because natural language sentences in the title and the body are good clues to understand the whole card. At this point, we don't use any formal language and mark up language since describing such language is a hard task for nonprofessional people.

Fig.1(a) shows the components of the knowledge card. The knowledge card consists of "title" and "body" sentences and an "image".

Title: The title represents a summary of the knowledge card. This is very useful for both humans and computers to understand the whole card. The title in Fig.1(a) shows this card describes about the system called "POC TV".

Body: The body represents detail of the knowledge card. This can be free text about the title of the card.

Image: The image represents a related figure, picture or movieclip. The image supplements the body text.

The knowledge card is represented by using XML form (Fig.1(b)). The $< opinion >$ element shows the whole card. The text in $< title >$ element shows the title, the text in $< body >$ element shows the body and the URI in $< img >$ element shows the location of image.

People can create their content by writing and arranging the knowledge cards. "Basic story" with "QA story" is the form of the content. By using the basic story form, people can give context to knowledge cards, each of which is context independent. The basic story consists of one title that summarize the whole story and a sequence of N knowledge cards that are arranged by order of topics. Fig.2(a) shows an example of the basic story that is about the system called "EgoChat". The first card (Card 1) introduces the story, the second card describes about the motivation of EgoChat, and the following cards explain details.

The QA story represents a conceivable conversation about the basic stories. By using QA story, people can create conversational content. The QA story consists of a question card that is a conceivable question and a sequence of answer cards that is a basic story answering the corresponding question. Fig.2(b) shows an example of the QA story. In Fig.2(b), the conceivable question is "What is the virtualized-ego?" and the following N answer cards answer it.

To manage knowledge cards mentioned above, we have developed EgoChatIII system (Kubota, 2002). The aim of EgoChatIII is to extend opportunity for people to present their content and get

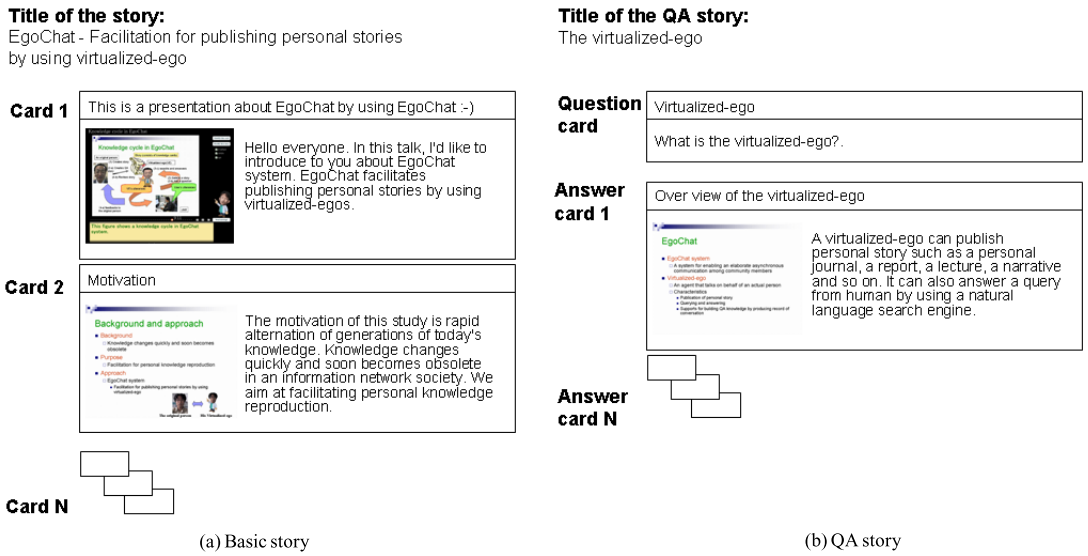


Figure 2: Examples of the story

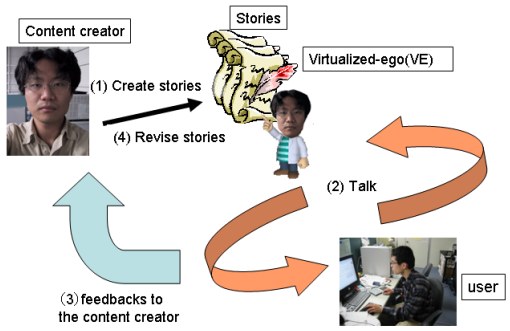


Figure 3: The content circulation mediated by the virtualized-ego

comments about them by using conversational agents. It is important for people to gather opinions about their content because the sustainable growth of the content needs a lot of comments. In EgoChatIII system, people can delegate their own agent called virtualized-ego to talk about their content.

The virtualized-ego circulates content between a content creator and a user as follows (Fig.3):

Process 1: The creator accumulates his content into his virtualized-ego by creating basic stories and QA stories by hand.

Process 2: The virtualized-ego can make a presentation to a user by reading out the body text and displaying the images in the basic stories. The virtualized-ego can also talk with the user by using full text search (Kiyota, 2002) from QA stories.

Process 3: The virtualized-ego brings user's opinions about the story back to its creator by email.

Process 4: The creator revises his stories referring to the user's opinion.

An example of the conversation using stories in Fig.2 is as follows:

Virtualized-ego: Hello everyone. In this talk, I'd like to introduce to you about EgoChat system. EgoChat facilitates publishing personal stories by using virtualized-egos.
User: What is the definition of the virtualized-ego?
Virtualized-ego: A virtualized-ego can publish personal story such as a personal journal, a report, a lecture, a narrative and so on...

The virtualized ego searches user's question from his question cards, then replies the answer cards corresponding to the question card. The full text search (Kiyota, 2002) allows a gap between user's question and a question card by resolving modifier-head relation gap [1] .

There are three advantages to using knowledge cards and : First, the creator needs not to care about formalizing content except boundaries of cards. Free text can be included in the card. Second, the knowledge card improves reusability of the content. It is easy for the creator to make variations of content because a constituent part of the content is represented by an independent card that can be briefly attached and detached with the content. because recent technologies of NLP are good at text processing in a paragraph like the body text of a knowledge card.

3 Knowledge channel

In the content management, the conversational agent like the virtualized-ego seems to be fit for supporting rough sketching process because conversation is the style for informal communication. However, the editorial space like a white board is necessary in the process of putting rough sketches into shape because conversation is not accumulative but temporary. To integrate the conversational process and the editorial process, this paper proposes a knowledge channel model that is a model of an agent-based content management system.

The knowledge channel model consists of a conversational agent and a landscape. Scattered stories are assembled into streamed content called "knowledge channel" by a channel policy. The channel policy is a policy document that enables people to declare the intention of managing content. The channel policy controls miscellaneous basic stories under one common policy written by a person who manages the channel. The channel policy is a XML document that consists of four main elements: (a) "flowStrategy" element describes the arrangement of stories. (b) "outlineStrategy" element describes the policy for viewing stories in the editorial process. (c) "delegationStrategy" element describes the policy for making a presentation in the conversational process. (d) "accessStrategy" describes access permission for stories.

The landscape supports the editorial process. It enables a user to understand a set of content spatially. This is not just visualization of content, but externalization of user's understanding about content. The user has leadership in the process of mapping (e.g. categorizing, scaling or locating) content on the landscape while the computer has leadership in the visualization system. Specially, the characteristic of the landscape is growth of the map according to progress of user's understanding. For example, the landscape supports a novice user by not only visualizing a primary map, but hinting the step to following growth of the map. The landscape system should not be a push system but a proposal system. The system must not relocate content without user's permission because the user may be confused if his map of understanding is changed by others.

A user can speak and listen to stories by using a conversational agent while he can view and construct stories by using a landscape. The conversational agent is suitable for activating and stirring stories while the landscape is suitable for fixing and arranging them. People can brush up their stories by repeating such two styles of interaction like annealing steal.

4 Implementation

The knowledge channel model is materialized on a web-based system where the landscape is partially implemented as a dynamic program table system that visualizes stories as a TV program table according to a channel policy. Fig.4 shows overview of the knowledge channel system. The

[1] This is indeed Japanese full text search, the English example above is just an approximation.

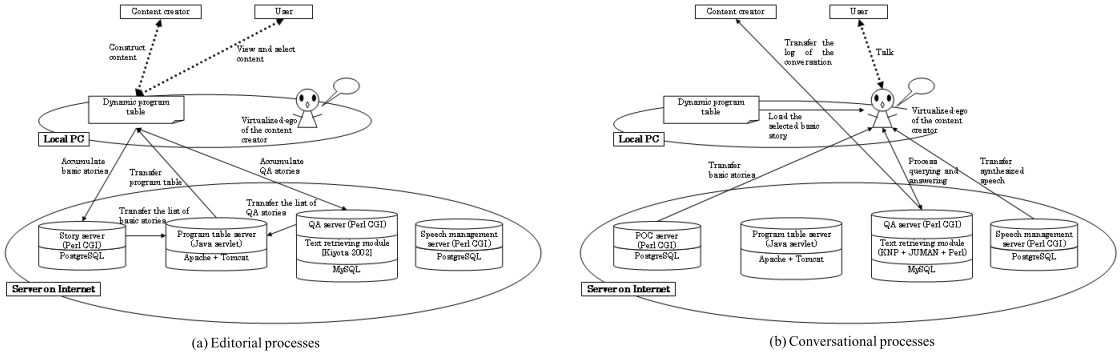


Figure 4: The knowledge channel system

knowledge channel consists of four servers (Story server, Program table server, QA server and Speech management server) and two web interfaces (Dynamic program table and virtualized-ego).

Story server: Story server accumulates basic stories with their creator, modified date and permission by using PostgreSQL. (This server is originally developed by Fukuhara(Fukuhara, 2001).)

Program table server: Program table server manages the arrangement of stories according to the channel policy.

QA server: QA server accumulates and retrieves QA stories by using text retrieving module (Kiyota, 2002). QA server also transfers the log of the conversation to the content creator.

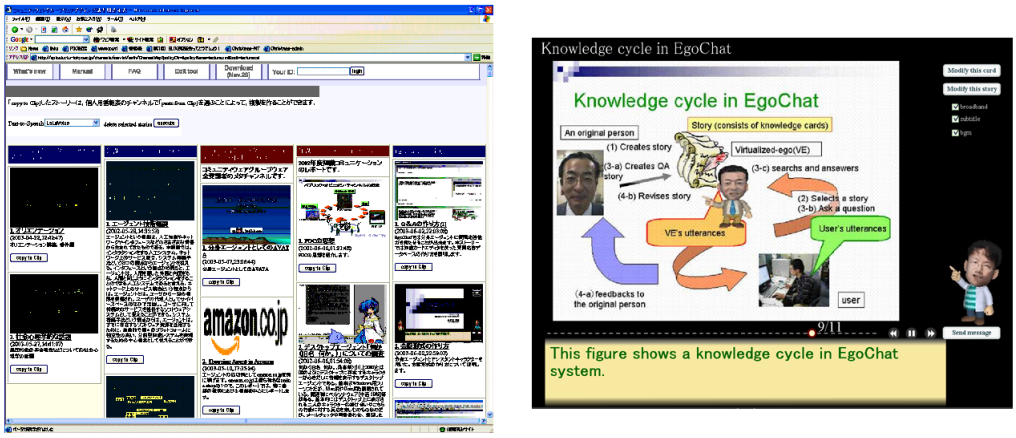
Speech management server: Speech management server manages real-time speech synthesis. This server provides CGI interface to speech synthesis systems and caches speeches that are once synthesized. In this implementation, we use Galatea talk [2] for Japanese speech synthesis and Festival [3] for English speech synthesis.

Dynamic program table: The dynamic program table is the web-based landscape that provides the overview of the stories in HTML form (Fig.5(a)). It also provides knowledge card editor developed by Dynamic HTML.

Virtualized-ego: The virtualized-ego is the web-based conversational agent that can talk about the content on behalf of the content creator. Fig.5(b) shows the screen of the virtualized-ego. The agent character with the face of the actual person is on the right end of the screen. It can read out the body text by using synthesized speech, show the image on the center of the screen and accept user's comment from the keyboard. It is implemented by Macromedia Flash.

Fig.4(a) shows the editorial processes where a content creator constructs his content and a user views and selects interested content on the dynamic program table. The dotted arrows represent human-computer interaction and the solid arrows represent the outline of the content flow. The dynamic program table shows the overview of stories on demand according to the list of basic stories and QA stories organized by the channel policy. The content creator can create and edit his stories by using the knowledge card editor on the dynamic program table, and then the content is accumulated into Story server and QA server. On the other hand, the user can view creator's content and select one of them on the dynamic program table, then the conversational processes start.

[2]http://hil.t.u-tokyo.ac.jp/ galatea/index.html
[3]http://www.cstr.ed.ac.uk/projects/festival/



(a) The screen of the dynamic program table

(b) The screen of the EgoChat

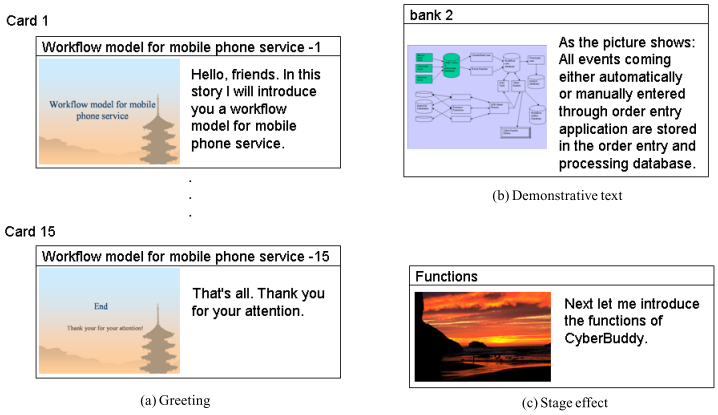
Figure 5: The screen shot of the knowledge channel system

Fig.4(b) shows the conversational processes where the virtualized-ego of the content creator talks with the user about the selected content. The user can ask and listen to the basic story that is read out by the virtualized-ego. The conversation between the user and the virtualized-ego are recorded and sent to the content creator.

5 Experiment

The knowledge channel has been experimentally used in 3 communities; "Lecture community", "RISTEX community" and "Conversational informatics community". We have provided 83 program tables that visualize 290 channels including 475 stories. "Lecture community" consists of a professor and class members. This community was formed together with the lecture of Prof. Nishida in Tokyo University (the first term: 2002/5/7 - 2002/8/9, the second term: 2003/4/8 - 2003/9/16). Two channels called "Knowledge communication (2002)" and "Knowledge communication - Paper" were managed in the first term. "Knowledge communication (2002)" is a channel about knowledge communication, which includes 27 stories (274 knowledge cards) created by the professor. "Knowledge communication - Paper" is a channel for papers about agent technologies and work-flow models, which includes 140 stories (1031 knowledge cards) submitted by 58 class members. In the second term, "Community-ware and groupware design (2003)" and "Community-ware and groupware design - Paper" were added to this community. "Community-ware and groupware design (2003)" includes an orientation story for the class members and a story about psychological issues of group behavior. "Community-ware and groupware design - Paper" is the channel for papers about agent technologies, awareness and community support systems, which includes 94 stories (988 knowledge cards) submitted by 43 class members.

Each class member in 2003 has a personal program table that integrates and visualizes three channels; "Personal channel" including personal secret content, "Public channel" including public open content and "Community-ware groupware channel" including his papers. The professor has a personal program table and a program table for grading. In the grading table, "Community-ware groupware channel" of the whole class members are arranged. The professor and the class members shared one program table called "'Community-ware and groupware' program table" that integrated and visualized four channels; "Knowledge communication (2002)", "Knowledge communication - Paper", "Community-ware and groupware design (2003)" and "Community-ware and groupware design - Paper". "Community-ware and groupware design - Paper" is a meta-channel that enables people to view plural channels as one channel by integrating 43 "Community-ware groupware channel"s.



(a) Greeting (b) Demonstrative text (c) Stage effect

Figure 6: Examples of the heterogeneous knowledge card

"RISTEX community" consists of researchers in RISTEX Mission-Oriented Research Program I [4] that aims at creating new technological system in order to solve social problems related to safety and security. This community has been formed to open the research results to citizens from Dec. 2003. 14 research groups have their own channels and publish them on the public program tables. Program tables are distinguished by "accessStrategy" (public or members only) and language (Japanese or English). The public tables are opened to everyone for communicating widely while the private tables are closed for professional discussion. Most of stories are Japanese, and public stories are less than private stories because the content that is opened to unspecified people forces more attention to content creators than private content.

"Conversational informatics community" consists of researchers in the conversational informatics project (Apr. 2001 - Mar. 2006) aided by Japan Society for the Promotion of Science. This community was formed to discuss research issues and publish research results from Sep. 2003. This community has three symposium channels that virtualize real symposiums and workshops of the conversational informatics project. In "Conversational informatics channel", 2 days symposium on conversational informatics is roughly summarized in 15 stories (total play time is 4 hour) created by 15 presenters. And other channels include keynote lectures and an interim report of this project.

6 Discussion

The result of the experiment suggests that the knowledge channel can manage huge content. The half-year lectures can be summarized into about 260 stories, and 2 days symposium can be summarized into about 15 stories. Students' papers are shared in the class. Moreover, all the content is accumulated in reusable form by using knowledge cards.

The knowledge card allows heterogeneous content. The lecture content is written in a formal style on the contrary the papers of class members include a casual style like a talk show. Fig.6(a) shows an example of the story that starts and ends with greeting. Fig.6(b) shows an example of the demonstrative text that complements the picture about workflow for a bank system. Each constituent part (title, body and image) is ambiguous, however the whole card makes a sense. Fig.6(c) shows an example of the stage effects. The image represents sunset that is not related to the body text directly, however it represents a change of topic by having a break.

The above results implies that the knowledge channel allows the management of huge heterogeneous content.

There are some related works in the personal content management area. Weblog systems

[4] http://www.ohriki.t.u-tokyo.ac.jp/S-Tech/M1/eng/index.html

(Rebecca, 2002) like Movable Type [5] and TypePad [6] support for a user to publish his personal journal by using a card like form called an "entry". They also support communication between users by the trackback function that enables users to link each other's entries. However, each entry is too independent to make a large content. On other hand, our knowledge cards can be organized into a long story. The knowledge channel differs in that the weblog supports text-based communication while the knowledge channel supports human-agent conversation.

7 Conclusion

This paper described the agent-based content management system that allows huge heterogeneous content. The knowledge card that can wrap up heterogeneous content has been proposed. The knowledge channel that is a model of agent-based content management system has been proposed and implemented by using virtualized-ego and dynamic program table. The empirical experiment conducted in three communities has implied that the knowledge channel allows the management of huge heterogeneous content.

References

Fukuhara, T., Azechi, S., Fujihara, N., Matsumura, K., Kubota, H., Nishida, T. (2001). Public Opinion Channel: Facilitating Community Knowledge Circulation, Workshop on Community Knowledge, ECSCW2001.

Kiyota, Y., Kurohashi, S., Kido, F. (2002). "DialogNavigator": A Questions Answering System based on Large Text Knowledge Base, Proceedings of The 19th International Conference on Computational Linguistics (COLING 2002), pp.460-466.

Kubota, H., Kurohashi, S., Nishida, T. (2002). Virtualized-egos using Knowledge Cards. In Proceedings of Seventh Pacific Rim International Conference on Artificial Intelligence(PRICAI-02) WS-5 International Workshop on Intelligent Media Technology for Communicative Reality (IMTCR2002), pages 51-54.

Kurohashi, S., Nagao, M. (1994). KN Parser : Japanese Dependency/Case Structure Analyzer. In Proceedings of The International Workshop on Sharable Natural Language Resources, pages 48-55.

Kurohashi, S., Nakamura, T., Matsumoto, Y., Nagao, M. (1994). Improvements of Japanese Morphological Analyzer JUMAN. In Proceedings of The International Workshop on Sharable Natural Language Resources, pages 22-28.

Rebecca, Blood (2002). "THE WEBLOG HANDBOOK", PerseusPublishing.

[5] http://www.movabletype.org/
[6] http://www.typepad.com/

The Measures for the Evaluation of Communication Tools: the causality between the intention and users' subjective estimation of community

Ken'ichi Matsumura
Research Institute of Science and Technology for Society, JST
ATAGO greenhills MORI Tower 18F,
Atago 2-5-1, Minato-ku,
Tokyo, 105-6218, Japan
matumura@ristex.jst.go.jp

Abstract

In this paper, we propose standardized measures for evaluation of communication tools, and present the method of evaluation for the community formed through the usage of communication tools. In the experiment of a test-bed community, the intention to active participant was influenced by benefits of tools, which community members felt. The intention to use tools continuously was influenced by their interests in the tools. The benefits and interests were affected by users' evaluation of their community. Users' subjective estimation of community consists of three factors; understanding the state of community, own contribution to their community, and clearness of others' opinion. These factors are expected as the measures for evaluation of communication tools for the community.

Keywords: Public Opinion Channel, SIQ, intention, benefit

1 Introduction

The users of the Internet are increasing now, and the way to use it become diverse. Some of them use the Internet in order to acquire information which they need, and others do for the purpose of sending message to general public. Along with users' increasing and diversity of purposes of usage, many web applications for communication are developed and investigated, for example, mailing list, BBS (Bulletin Board System), and so on. But we don't have the standardized measures of evaluation for communication tools. How should we evaluate these communication tools? By developing standardized measures, we can easily compare between communication tools. This comparison enables us to find characteristics of tools and difference from other communication tools clearer. Additionally, results of evaluation by standardized measures will point out improvement points in a communication tool. After improvement, re-evaluation using the same measures will show the effect of improvement. Therefore, the standardized measures of evaluation for communication tools are necessary to compare with them.

In this paper, we discuss the factors which form users' intention to participate in a community, and the method of evaluation of the communication tools from users' psychological perspective through the experiment of Public Opinion Channel.

2 The Aproach for Evaluation of Public Opinion Channel

2.1 Public Opinion Channel

The Public Opinion Channel (POC) is an automatic community broadcasting system that elicits opinions from community members, generates a story from them, and broadcasts the story to the community (Azechi et al, 2001). Fig.1 shows the concept of POC. Our prototype POC system

consists of a community broadcasting servers (POCservers) that broadcast opinions to POC clients that automatically display them.

The difference from POC and BBS is the way of displaying messages contributed to the community. In spite of POC, messages contributed by community members are repeated to display automatically. Community members can select messages what they wanted. Therefore members of POC community are not likely to be influenced by context. On the other hand, members of BBS community need to operate system for watching messages.

In the case of POC, the benefit of using it consists of easiness for sending and acquiring information, which are important points to evaluate.

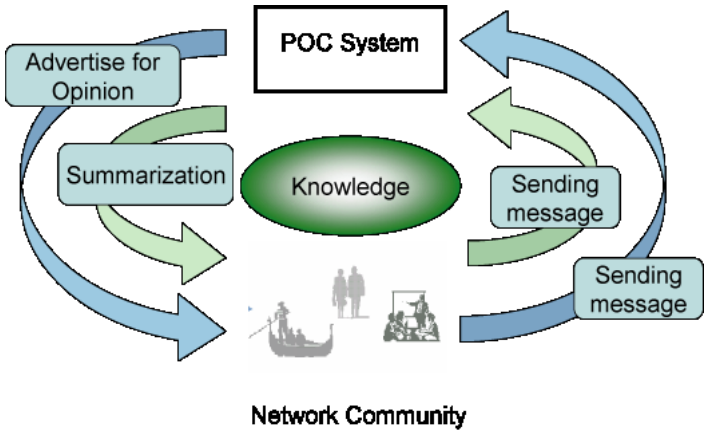


Figure 1: The concept of Public Opinion Channel

2.2 SIQ

Yamashita and Nishida (2002) proposed a concept of evaluation framework communication systems called Social Intelligence Quantity (SIQ). SIQ consists of SIQ-Personal and SIQ-Collective. SIQ-Personal is measures for users' psychological perspectives. On the other hand, SIQ-Collective is measured for community's perspective like amount of information in a community, and objective indices such as diversity and convergence of information. In this paper, we focus on SIQ-Personal, and discuss the intention to use a communication tool continually, and to participate in a community actively. Specifically, we introduce the trial of evaluation for POC community members' subjective estimation to their community.

Many indices are considered as measure for the evaluation of community, the number of access, the amount of information, accumulated information, and so on. Fukuhara et al., (2003) developed POCanalyzer, which shows mainly the number of access to the community, and the number of information sent by community members. Although these data show a part of members' behavior in the community, analysis with users' psychological side make it possible to evaluate more strictly for a communication tool.

2.2.1 The intention

Community members have various purposes. For example some members may search information on good restaurants, and others may look anybody who they talk with.

Fishbein and Ajzen (1975) presented the Theory of Reasoned Action (TRA). In this theory, the behavioural intention is influenced by attitude to behaviour and subjective social norm. Behavioural intention is the representation about the action exacted in the future, see Bandula (2001). TRA assumes that people act on the base of behavioural intention formed. Therefore, measurement of behavioural intention is beneficial to predict future behaviour of users. And understanding users' intention will make it possible to indicate that the problem. For example, when

user with high intention doesn't act on, we can guess that there might be some factors preventing users' action outside of communication tools.

If there exist these factors in the community, they sometimes don't act on their intention. On the contrary, if users' intention will not be formed, we can't guess that a tool could motivate them to use it. Therefore, the measurement of behavioural intention will make it possible to focus the problem of the tool. And indicate point of improvement. Davis et al (1989) indicated that TRA can be applied to acceptant of technology in an organization. Their model predicts acceptance and usage of information systems in an organization. In this model, behavioural intention is influenced by perceived utility, perceived easily usage, and subjective social norm. Subjective social norm is defined as perceived expectation from other members. Though subjective social norm influence the intention in an organization, it has few effects in a network community, because there are few hints to judge social norm in network community. Therefore, in this study, we discuss the relation among behavioural intention, users' subjective evaluation of the POC system, and their subjective evaluation of their community. In this study, we discuss the received benefit of using a tool and the interest as subjective evaluation of POC system. And understanding the state of a community, the clearness of others' opinion, and the contribution to a community are discussed as the subjective evaluation of a community.

2.2.2 Benefits of using a tool and Interest in a tool

Benefits of using a tool include both sending and acquiring information. If users can find benefits of using a tool, they would have the intention to continue to use them, and to participate actively in a community. On the contrary, they can't find benefits of using a tool, they would neither want to use them nor join that community. Of course, if users are not interested in a communication tool, they wouldn't try to use them. Therefore, the benefit and the interest are considered as influential factors on the intention. Next, we discuss factors which are predicted have an effect on the benefits of using a tool.

2.2.3 Understanding the state of community

Understanding the state of a community includes follows: understanding others' style of behavior, acceptance of reply, the interaction among community members. It is very important for users to understand the state of community and others' behavior on judging own contribution to their community. And understanding the state of community will make it easy for user to judge activeness of communication in the community.

On the contrary, if they couldn't understand the state of community, they can't judge the activity of communication in a community.

2.2.4 Clearness of others' thought

Understanding the state of community includes the interaction with other members, in contrast the clearness of others' opinions is whether community members can see contents from other members. If community members can see others' opinion, they can find the information which they want to acquire. Additionally, the clearness will make it possible to compare own opinions with others' them. This lead to add own information with lacking information. Therefore, the clearness of others' opinion has an effect on the benefit of a tool which they feel.

2.2.5 Contribution to their community

The contribution includes community members sending information and answering to question. In the situation that community members can understand both of the state of community and others' opinion, they could identify their own positions and roles in a community. The contribution to their community may make members feel self-efficacy, and motivate them to participate actively.

3 Method

Participants were 50 members living in Kansai area in Japan. They were divided into 5 communities. Ten participants of them joined a community and asked to participate in the community for 4 weeks. All participants were asked to join the POC community by using POCviewer on the internet. They talked about domestic topics in the Kansai area of Japan. Starting members continued to talk about that after new comers joined the POC community.

They were asked to answer questionnaire about the intention, subjective evaluation of tools, and the benefits of using POC system after 2weeks usage. They had to send 14 messages to their community in four weeks.

4 Results

Each factor was measured in 1 (negative) to 5 (positive). Table 1 shows means and SD of scores of each factor. All factors are estimated positively. Particularly, the interest was estimated very positively. We did covariance structure analysis in order to examine the causality between the intention and subjective estimation. Figure 2 shows the causalities of intention and factors which we predicted influence on the intention.

Table 1: Means of scores on each factor

	Means	SD
Intention to paricipate actively	3.64	0.94
Intention to use tools for long-term	3.80	0.83
Benefit of tools	3.85	0.88
Interest in tools	4.05	0.83
Contribution	3.20	0.89
Clearness	3.54	0.85
Understanding	3.46	0.73

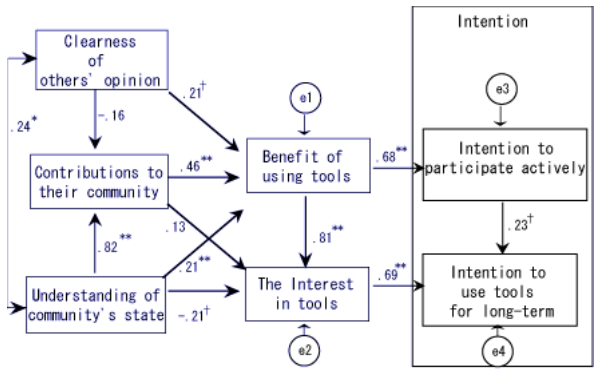


Figure 2: The model of intention to participate in the community

4.1 The effect of the benefit and interest on the intention

The intention to participate actively was influenced strongly by benefits of using POC, and the intention to use POC for long-term was influenced by the interest in tools, which was affected by benefit of using tools.

In other words, when users recognize that the tool has benefit of using, they become to have an intention to participate in the community formed by POC, and are interested in POC. The interest in POC heighten the intention to use POC for long-term.

Therefore, we can say that the subjective evaluation to POC have an important role to heighten the intention to participate in the community.

4.2 The effect of Clearness, understanding and contribution

The understanding of activity in the community has a positive effect on the benefit, but also that has a negative effect on their interest in the communication tools. On the contrary, the benefit influenced on the interest, the understanding of community's state had an indirect effect on the interest through the benefit of tools.

The understanding affected on the contribution strongly. They, who felt that they could contribute to other community members, tended to evaluate the communication tool had much benefit.

The clearness of others' opinion had an effect on the benefit of tool. In the case that they could see others' opinion clearly, that would make it easy for users to make choices of information which they needed. Therefore, the clearness has an effect on the perceived benefit.

The contribution was influenced by the understanding and the clearness, and had effects on both the perceived benefit and the interest in POC.

In other words, community members can understand other members' activities and others' opinions in POC community, and their understanding activities and ideas made users feel the contribution to the community. These understanding and the subjective contribution had a positive effect on the perceived benefit of using tools.

5 Discussion & Future work

These results indicated that the intention to participate in a community was influenced by the received benefit of using a communication tool. And the benefit was affected by factors associated with understanding the state of a community. Particularly, users' subjective contribution had a strong effect on the benefits of using a tool. This means that whether users can contribute to the community affects the benefits which they find. As these benefits affected the intention to participate actively, it is important for users to feel contribution in their community. The contribution reflected the understanding the state of a community and the clearness of others' opinion, and had a positive effect on the benefits of using a tool. Therefore, in order to evaluate a communication tool for supporting a community, we should examine not only users' subjective evaluation of a communication tool but also their subjective evaluation of community.

These results indicate that we need to consider factors which have effects on their intention so that users' intentions are measured as an index of SIQ-personal. In order to evaluate communication tools more strictly, we need (1) to examine the relation of members' subjective evaluation and indices of their activities, and (2) examine the effect of quantity and quality of information stored in the community.

In this study, the intention to participate in the community was formed by the benefit, which was affected by the subjective contribution to the community and understanding the state of a community strongly. Additionally, if the intention doesn't lead to action, we need to specify factor which prevent their activities. We need to construct the method of standardized measures as like SIQ in consideration of these factors, in order to point out improvement through evaluation examination. In the future, we need to examine the relation of members' intention and indices of their activities in communities formed by other communication tools in order to construct standardized measures for evaluation of communication tools.

6 Conclusion

In this paper, we discussed measures for evaluation of a communication tool from personal psychological perspectives. In POC community, the intention to participate in a community was

influenced by the received benefit of a tool, and understanding the state of a community and the contribution to their community had effects on the benefit. These results indicated that we need to consider both subjective estimation of a tool and a community in order to examine whether the intention to participate in a community are formed by using a communication tool.

So that means, in evaluation of communication tools, we need to consider these factors from users' perspectives. It will be possible to evaluate communication tools more valuably by consideration of these factors. In the future, we need to examine the relation of members' intention and indices of their activities in communities formed by other communication tools in order to construct standardized measures for evaluation of communication tools.

References

Azechi, S., Fukuhara, T., Fujihara, N., Sumi, K., Matsumura, K., Hirata, T., Yano, H., Nishida, T. (2001). "Public Opinion Channel–Toward Knowledge-Creating Community-": *Journal of-Japanese Society for Artificial Intelligence*, Vol.16, No.1,pp130-138, 2001

Bandula, A. (2001). Social Cognitive Theory: An Agentic Perspective, In Annual Review of Psychology, Vol. 52: 1-26, 2001

Davis, F. D., Baggozzi, R. P., and Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models, Management Science, Vol.35, 8, 982-1003

Fishbein, M., and Ajzen I. (1975). Belief, attitude, intention and behavior. In An introduction to theory and research

Fukuhara, T., Chikama, M. and Nishida, T. (2003). A Platform for Investigating a Knowledge-Creating Community: Community Analysis and Maintenance Functions in the Public Opinion Channel, In Social Intelligence Design International Conference 2003

Yamashita, K. and Nishida, T. (2002). SIQ(Social Intelligence Quantity): Evaluation Package for Network Communication tools. In The 5th Asia Pacific Conference on Computer Human Interaction

Retrieving information on the World Wide Web: Effects of domain specific knowledge

Asako Miura
Kobe Gakuin University
518 Arise Ikawadani-cho, Nishi, Kobe, 651-2180, Japan
asarin@team1mile.com

Nobuhiko Fujihara
Naruto University of Education
748, Nakajima, Takashima, Naruto-cho, Naruto, Tokushima, 772-8502 Japan
fujihara@naruto-u.ac.jp

Koji Yamashita
National Institute of Information and Communications Technology
3-5 Hikaridai, Seika, Soraku, Kyoto, 619-0289 Japan
koji@nict.go.jp

Abstract

In this study, we intend to examine information retrieval behaviors from a psychological point of view using a search engine on the World Wide Web (WWW). We investigated information retrieving behaviors in detail based on both recorded data of retrievers' web browsing actions and their thinking processes by the "think aloud" method. We focused on selected keywords for retrieving and compared them between retrievers who had enough knowledge about their task and those who did not. Our goal was to learn about the literacy of finding required information efficiently on the WWW.

Keywords: Information Retrieval, Search Engine, Domain Specific Knowledge, Think Aloud.

1 Information retrieval on the WWW

Information and network technologies are being applied to every aspect of our everyday life. An enormous amount of information has been accumulated in electronic media by some huge projects such as e-library and the information range searchable for us has been extended more than ever before. Especially after the appearance of the World Wide Web (WWW) in 1992, patterns of information distribution via the Internet changed completely. The WWW enables us to transmit privately any information we had to everyone logging on the Internet. As a result, the Internet turned out to be a medium for not only collecting but also transmitting information for us. In other words, especially on the Internet, the quality of information itself drastically changed its characteristics from "classified and organized" to "unclassified and amorphous." Though the WWW has become, a garbage dump of poor or incorrect information as well as a gold mine of valuable information, it can be a useful tool for knowledge acquisition and creation depending on how we use it. It is necessary for Internet users to discriminate truth from falsehood and find their needed information efficiently (i.e., rapidly and exactly).

When we search for information on the WWW, we usually try to retrieve it by using search engines such as Google. For rapid and exact information retrieval, we have to attain certain literacy about how to treat information that we have acquired as well as a skill for how to use these search

engines. To describe such literacy concretely, information searchers on the WWW always need to pay attention to collecting sufficient information and only extracting information necessary to them from enormous retrieval results that include results irrelevant to them. In traditional studies of information retrieval systems, there are two standard measures of performance. The first is precision that is defined as the number of relevant documents retrieved divided by the total number of documents retrieved. The second is recall that is defined as the number of relevant documents retrieved divided by the total number of relevant documents in the collection. A system attempts to maximize both recall and precision simultaneously. If we apply these standards to human behavior, precision would be to collect sufficient information and recall would be to extracting necessary information. Recent research on searching behavior on the WWW using search engines (Kim (2001); Kim and Allen (2002); Palmquist and Kim (2000); Wang et al (2000); Wolfram and Dimitroff (1998)) has also shown that people's WWW search behaviors are different from traditional information retrieval patterns. For considering such literacy from a psychological perspective or developing a supporting system for information retrieval on the WWW, we have to examine retrieving behaviors thoroughly and verify how they associate with efficient/inefficient search results.

2 Our previous studies

In considering information searching behavior from a psychological point of view, it can be regarded as a perfect example of a decision-making or problem-solving process in a highly-networked information society. It is particularly interesting that there seems to be some dynamic interactions between pre-existed knowledge of searchers and external information of retrieved results by search engines. In such an interaction, information retrievers may find some unexpected novel knowledge as well as acquire information they have needed.

Table 1 Categories and subcategories of retrieval actions

1. Selection of keywords	
1-1	Input of keywords
1-2	Add keywords
1-3	Delete keywords
1-4	Substitute keywords
2. Operation of web pages	
2-1	Select hyperlinks
2-2	Select item of pull-down menu
3. Operation to browser	
3-1	Press forward button
3-2	Press back button
3-3	Press home button
3-4	Select URL from jump menu
3-5	Input URL directly
3-6	Select other windows
4. Miscellaneous	

Our previous studies, Fujihara and Miura (2003) and Miura and Fujihara (2001), carried out a psychological experiment on information retrieval and investigated the retrieval process of average Internet users. We analyzed exhaustively all kinds of retrievers' actions and classified them into some categories. In these studies, graduates and undergraduates were asked to retrieve information for two tasks using a search engine. They retrieved information on the WWW regularly but were not so specialized about that. One task was well-defined and had only one answer and the other was ill-defined and had several possible answers. There was no difference in task-related knowledge among the participants. Actions for information retrieval by search engine were categorized into four categories: selection of keywords, operation on web pages, operation on browser, and miscellaneous (detailed definition and subcategories are shown in Table 1). Result

revealed that the average of the number of actions for an ill-defined task was larger than that for a well-defined task ($p < .05$). On the other hand, there were no significant differences for the number of web pages browsed, the search times between tasks, the frequencies of appearance of each action category, and the number of keywords used for retrieval. These results suggest that average Internet users have only a simple strategy and apply it to any task for information retrieval on the WWW. In this way, some basic actions for information retrieval on the WWW were categorized and we were able to estimate various retrieval behaviors based on these categories. There was no difference among tasks when participants have nearly equal task-related knowledge and skill for information retrieval.

According to these studies, based on these categorized retrieval actions, we compared information retrieval behavior between those who have enough task-related knowledge and those who do not. As it can be predicted that task-related knowledge would have some effects on the process of selecting keywords for information retrieval, we focus on how an effective information retriever behaves.

3 Experiment

3.1 Participants

Twelve graduates and undergraduates of Osaka University and Naruto University of Education (8 female and 4 males, mean age=21.33 years) participated in this experiment. Six of them were psychology majors and the remaining six were education majors. As experimental tasks described later were associated with psychology, the former were assumed to be "knowledgeable" participants and the latter were "unknowledgeable." Some attributes of each group are summarized in Table 2. Table 2 indicates that the participants in the knowledgeable group had used a PC and the Internet longer and made informational retrievals on the WWW more frequently than those in the unknowledgeable group. Though we have to consider this difference when we investigate our results, we also need to take into consideration that high-frequency informational retrievals do not necessarily correspond to their level of proficiency directly. Some rewards were given for their participation.

Table 2 Attributes of participants

	Knowledgable	Unknowledgeable
N. of participants	6	6
Age	21.67	21.00
Years of computer use	4.67	3.33
Years of Internet use	4.67	3.33
Hours of Internet use a week	8.50	2.17
Frequency of web retrieval a week (persons)		
almost every day	3	0
4-5 days a week	1	0
2-3 days a week	1	4
1 day a week	1	2

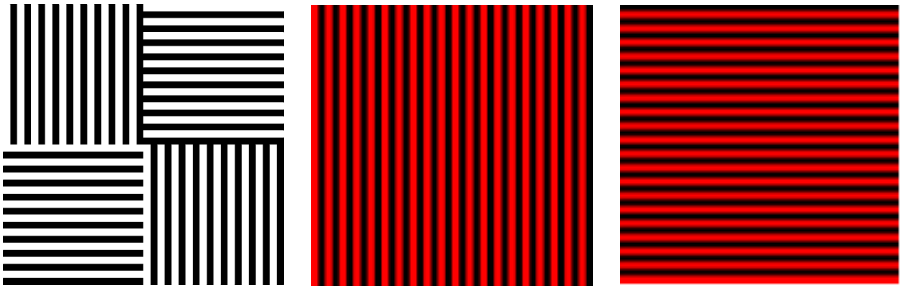
3.2 Tasks

The participants were asked to search for web pages to perform the following two tasks (named task A and B) and to provide the URLs of the pages that provided the answers. Task A was associated with perceptual psychology and task B was associated with social psychology. The psychological phenomena introduced in each of the tasks were easy to understand but the psychological technical term that represents the phenomena must not be generally known.

- Task A: This task was related to one of the color aftereffects in human visual perception. Before beginning information retrieval, the participants were asked to gaze at stimulus figures (one test stimulus and two adaptation stimuli shown in Figure 1) for two minutes and they experienced this color aftereffect. After that, they were asked to retrieve the term by which

this effect was named, after the name of the researcher, by using a search engine on the WWW. The answer was "McCollough Effect."

- Task B: This task was related to one of the social influence phenomena, or social loafing, in human group behavior. Before beginning information retrieval, the participants watched a motion picture that explained social loafing with a tug of war. After that, they were asked to retrieve the term by which this effect was named, after the name of the researcher, by using a search engine on the WWW. The answer was "Ringelman Effect."



Test Stimulus Adaption Stimuli (left: red x black; right: green x black)

Figure 1: Stimulus figures for Task A

3.3 Procedure

The participants conducted information retrievals using a notebook PC (SONY VAIO; Windows XP OS). The PC was connected to the Internet through a LAN. They were directed to vocalize all of their thoughts and passing ideas (i.e. the think aloud method) through their information retrieval. Google, which can be regarded as the most familiar search engine in Japan, was selected as the retrieval tool in this experiment. To give them time to adjust to thinking aloud, the participants performed a trial information retrieval session. The trial task was "What is the longest river in Hokkaido, Japan and how long is that?" and the answer is "Ishikari-river, 268 km long."

The participants were asked to answer the question and provide the URL that contained the answer within 20 minutes. If they had not found the answer by then, the retrieval session was forced to terminate. The participants performed both tasks and the order of solving the tasks was counterbalanced. If some participants could not think aloud sufficiently or clearly, the experimenter called their attention to thinking aloud during the session. After the end of the task, participants were asked to answer some questions about their previous knowledge of the task and the degree of confidence in their answer. After finishing the first task, they were told the other task and asked to answer it in the same way as the first task. To estimate the participants' retrieval performance, we applied the following two recorded data. The first was a sequential motion picture of retrieval actions that was recorded with screen capturing software named "Koushi-Daiko XP (by Lifetree Corp.)." The second was audio data of thinking aloud that was recorded by microphone or IC recorder. After finishing the two tasks, participants were asked questions on their demographic attributes and experience regarding computer use, Internet use, and search engine use.

4 Results

In this article, we focused on the detailed analysis of Task A retrieval performance. According to the post-survey, it was demonstrated that no participant had known the correct answer of the task

before their information retrieval in this experiment. There were 6 of 6 knowledgeable participants and 2 of 6 unknowledgeable participants who could obtain the correct answer.

4.1 Categorization of retrieval actions

Based on a preliminary examination of the screen-captured data, information retrieval behavior each participant was described as a series of actions and those actions were divided into four categories determined in our previous study (Miura and Fujihara (2001)). The number of web pages browsed, the search times, and the number of actions taken by each participant are shown in Table 3. All of the knowledgeable participants found the correct answer. Half of them did it in the shortest time and the number of their actions and keywords were also very small. These three participants seemed to obtain positive results with nearly minimum efforts. On the other hand, the unknowledgeable participants took a fairly long time and numerous actions for information retrieval even if they got the correct answer. They seem to repeat trial and error many times and check around various contents of the WWW during their retrieval session (and not a few participants' activities were in vain).

Table 3 Retrieval performance of each participant

Knowledgeable group						
Participants	K1	K2	K3	K4	K5	K K6
N. of browsed web pages	24	6	8	9	13	34
Times for searches (min)	19m43s	1m52s	3m26s	2m18s	12m48s	15m48s
Answer[1]	CO	CO	CO	CO	CO	CO
N. of actions	75	10	17	16	53	64
N. of keywords-related actions	18	1	3	3	15	11
N. of input keywords	17	1	3	3	13	9
N. if input keywords a kr-action	5.11	1.00	2.00	2.67	2.13	2.18

Unknowledgeable group						
Participants	U1	U2	U3	U4	U5	U U6
N. of browsed web pages	47	8	17	22	91	53
Times for searches (min)[2]	TO	8m4s	15m0s	TO	TO	21m10s
Answer[3]	NO	CO	CO	NO	NO	FA
N. of actions	122	20	51	63	218	103
N. of keywords-related actions	20	6	19	14	24	5
N. of input keywords	14	7	16	18	15	7
N. if input keywords a kr-action	1.60	3.67	2.42	2.86	2.00	2.80

The average rates at which actions of each category appeared among those with the subject attributes knowledgeable and unknowledgeable are shown in Table 4. Whereas, there was a distinct difference of quantitative retrieval performance between the groups, the average frequency for each action category was not so different between the groups. In previous studies, we found that there was no significant difference in the frequencies of appearance of each action category. Furthermore, in this study, it was suggested that there was no difference between knowledgeable and unknowledgeable subjects in a general strategy for information retrieval on the WWW.

[1]The participant finshed the task with CO: correct answer
[2]TO: Time over
[3]The participant finished the task with NO: No answer, CO: correct answer and FA: false answer

Table 4. Average frequency for which actions in each category appeared (%)

	Knowledgeable	Unknowledgeable
1. Selection of keywords	0.22	0.15
2. Operation of webpages	0.41	0.42
3. Operation of browser	0.38	0.42
4. Miscellaneous	0.00	0.00

4.2 Input Keyword

The total number of keywords input by each participant (some of them were repeatedly input) is also shown in Table 3.

In the next phase, these input keywords were classified into two categories depending on their contexts and meanings. The first criterion for classification was based on the information source of each keyword. We divided the keywords into one of 4 categories, 1) instruction-related: source from the instruction of the experimenter (e.g., green, pinstripe, and phenomenon), 2) search-results-related: source from certain words appearing during retrieval session, 3) inference/pre-existing knowledge: source in inference or pre-existed knowledge of each participant, and 4) miscellaneous. As there were some compound words which were combined multiple nouns, they were divided into single nouns and classified separately. The second criterion was based on the meaning of each keyword. We also divided the keywords into 1 of 4 categories, 1) description: directly related to the phenomenon participants had observed (e.g., red and pinstripe), 2) domain specific knowledge: academic term related to the phenomenon (e.g., visual illusion, color aftereffect), 3) mistaken notion: approximate but unrelated to the phenomenon (e.g., Fechner, entropy), and 4) unrelated to the phenomenon but possible cue of retrieval (e.g., researcher, effect, and experiment). The result of the categorization of the input keywords by the two criteria is shown in Table 5.

Table 5. Categorization for the input keywords on the two criteria

	Knowledgeable	Unknowledgeable
Criteria 1 : Information Sources		
instruction-related	18(0.37)	39(0.41)
search results-related	13(0.27)	26(0.28)
inference/pre-existing knowledge	18(0.37)	29(0.31)
miscellaneous	0(0.00)	0(0.00)
Criteria 2 : Meaning		
description	21(0.46)	29(0.38)
domain specific knowledge	13(0.28)	17(0.22)
mistaken notion	3(0.06)	14(0.18)
possible cue of retrieval	9(0.19)	17(0.22)

Comparing the criterion of information sources between groups, the participants in the knowledgeable group tended to more frequently introduce their inference or pre-existed knowledge into their retrieval. It was also characteristic for the unknowledgeable group that compound words were used as keywords (13 for unknowledgeable vs. 2 for knowledgeable).

Comparing the criterion of meaning between groups, the participants in the knowledgeable group tended to use domain-related keywords (description and domain specific knowledge) more frequently than those of the unknowledgeable group. Especially, the participants who found the correct answer in the shortest time mainly used domain specific keywords from the start of retrieval. On the other hand, the frequency of mistaken notion in the unknowledgeable group was relatively larger than that of the knowledgeable group.

4.3 Protocol analysis of think-aloud performance

The think aloud performance of each of the 12 participants was transcribed into protocol data. We give a rough outline of comparison of those protocol data between knowledgeable and unknowledgeable groups. In the knowledgeable group, participants tended to find and focus attention on

some academic terms that represented the phenomenon they observed (e.g. aftereffect and complementary color) and immediately adopted them as keywords in the process of their retrieval. They also tended to read the contents of web pages linked to search results very carefully and tried to judge whether they were associated with the task or not. In cases of the unknowledgeable group, the participants tended to have difficulties specifying the certain domain of the given task and blindly focused on some proper names like researcher names and technical terms without considering their validity for the task.

5 Discussion

5.1 Summary of results

In this experiment, based on various indices of behavioral performance data, we compared information retrieval behavior between those who have enough task-related knowledge and those who do not. Results suggest that task-related knowledge had a great impact on the process of information retrieval in many stages. In this chapter, according to these results, we first consider some effective strategies for information retrieval in particulars based on the relationship between retrieval behavior and pre-existing knowledge. Second, we intend to advance some suggestions about future systems and another new perspective for information retrieval.

5.2 Relationship between retrieval behavior and pre-existing knowledge

Summing up our analysis of performance data in the previous and in this experimental study, there are two kinds of knowledge we should use for accurate information retrieval, 1) task-related domain specific knowledge, 2) knowledge (or skill) relevant to search engines or web browsing.

Whether participants have some task-related domain specific knowledge would have a much greater impact on various stages of their retrieval behavior. At the first stage of retrieval, if they have no domain specific knowledge relevant to a certain task, they immediately face some difficulties. They might not be able to comprehend which domain the task refers to (e.g. psychology in the case of Task A) nor select sufficient keyword(s) for filtering their retrieval results at first. In the next stage, browsing search results, domain specific knowledge also influences their retrieval behavior. In the default setting of Google (also in this experiment), search results include some significant cues with hyperlinks for retrieved URLs, such as the title of the web page found, text that is an excerpt from the returned result page showing your query terms in bold, and the cached link that enables retrievers to see the contents of the web page as of the time Google indexed it. If retrievers have enough task-related domain specific knowledge, they would make full use of these cues to filter the results. Lack of domain specific knowledge would lead to unproductive increment in performing their retrieval trials. Finally, domain specific knowledge has a great impact on the stage of assessing the likelihood of their answer. As the reliability of information on the WWW is quite ambiguous and uncertain, retrievers have to confirm it by themselves. Domain specific knowledge would make it easy for them to judge whether their answer is correct or incorrect.

According to the results of this study, it became clear that participants who had task-related domain specific knowledge (knowledgeable for psychology) performed more efficient (i.e., rapid and exact) information retrieval than those who did not. Furthermore, their retrieval performance tended to become more efficient when knowledgeable participants combine some technical keywords derived from their domain specific knowledge (e.g., visual illusion and accidental color) with keywords that directly related to the task (e.g., stripe and red). It would suggest that, through such combinations, they could effectively filter information from a vast array of search results.

5.3 Strategies and systems for information retrieval

These results suggest some efficient retrieval strategies in consideration of the relationship between the pre-existing knowledge of the retrievers and their given task. If retrievers have sufficient domain specific knowledge about a task, they should make positive efforts to use them as retrieval

keywords. That knowledge would make it possible for them to cleverly limit the scope of retrieval and extract more exact information while eliminating unnecessary details. If retrievers do not have sufficient domain specific knowledge about a task, especially when they have some difficulties in retrieving the correct answer with phenomenon-descriptive keywords, they should not expend much effort to get a definite answer, but firstly to comprehend a task-related domain for efficient retrieval. The latter knowledge that we should use for accurate information retrieval is associated with literacy about human information behavior. Compared with possible difficulties in acquiring sufficient domain specific knowledge, it would be easy for us to support information retrievers with some education or training programs and improvement systems or their interfaces.

As indicated in Fiske and Tayler (1999), however, a human being could be a cognitive miser, who usually tends to conserve their energy and to behave in a manner consistent with what was appropriate to gain useful information. Our results of previous studies and this study were strongly supported by the idea of a cognitive miser. Even if retrievers could not find their answer successfully, they seemed to be reluctant to change their retrieval strategies or try other actions than usual. They tended to run over "keyword search" and "web browsing" all in the same key. In sum, users of the WWW in general cannot make full use of search engines though they use them frequently.

When we intend to develop such a support system or interface that would enable us to retrieve information efficiently, it would be important to take into account the characteristics of the cognitive and behavioral aspects of human beings. It would be important not only to upgrade and expand the support system's functions but also to devise a user-friendly interface that would make active use of such functions. It might be effective to introduce social information filtering into retrieval results. Social information filtering is a general approach to personalized information filtering. Through social information filtering, items are recommended to a user based upon values assigned by other people with similar taste. The system determines which users have similar taste via standard formulas for computing statistical correlations. If we apply such a function in information retrieval systems, retrievers who do not have any pre-existing domain specific knowledge would be able to substitute recommended items with pre-existing knowledge. Alternatively, in cooperative retrieval, multiple members performing one common retrieval task together might promote the efficiency of information retrieval. It enables retrievers to supplement their lack of knowledge by exchanging their knowledge with each other. This kind of knowledge exchange would make it possible not only to compensate for the lack of knowledge of each retriever but also to have completely novel knowledge emerge through dynamic interaction between multiple members and in human-computer interaction (HCI). We will consider this possibility in the near future.

5.4 Limitation of this experiment

In this experiment, there was a particular limitation in collecting think aloud data. Though we asked all participants to vocalize their "thoughts" and "passing ideas", it seemed they did not necessarily follow our instructions faithfully. Judging from the results of the transcription, a considerable number of participants did not vocalize their thoughts or passing ideas but their "actions." Particularly in browsing retrieval results or linked sites they selected, participants seemed to be almost incapable of thinking aloud and there were few transcribed records. Though it could be surmised that participants in such phases thought or judged certain important considerations related to their retrieval, we were obliged to reduce our consideration of that.

There are some possible reasons for this limitation. Those were caused by the artificiality of participants performing "thinking aloud" during their retrieval session. One of the possible reasons was the brief degree of previous practice in thinking aloud and that might cause poor performance. In the next experiment, participants should be trained sufficiently prior to the experiment. Another possible reason could emerge from the shy characteristics of Japanese. In this experiment, the participants performed their tasks independently and the experimenter monitored their performance from behind. They might feel the weight of the eyes of the experimenter and not be able to think their ideas or thoughts aloud or adequately because of embarrassment. Corresponding to this possibility, it would be effective that we conduct an "ice-breaking" session

with them prior to the experiment or/and an interview as the occasion demands during the experiment. Anyway, based on the results of this study, further investigation into the details of the thinking process of retrievers is required for us to clarify human behavior and cognition in information retrieval on the WWW.

References

Fiske, S. and Taylor, S. (1999). Self-efficacy and personal control. In Fiske, S. and Taylor, S. (Eds.), *Social Cognition*, pages 197-204. McGraw Hill, New York.

Fujihara N. and Miura, A. (2003). Patterns of searching for information on the World Wide Web: A pilot study. *Psychological Reports*, 92, pages 1091-1096.

Kim, K.S. (2001). Information seeking on the web: Effects of user and task variables. *Library and Information Science Research*, 23, pages 233-255.

Kim, K.S. and Allen, B. (2002). Cognitive and task influences on web searching behavior. *Journal of the American Society for Information Science and Technology*, 53, pages 109-119.

Miura, A. and Fujihara, N. (2001) Experimental study of information searching behavior on www: Categorization of behaviors and examination of difference between tasks. *Transactions of Japanese Society for Information and Systems in Education*, 18, pages 121-128. (In Japanese with English abstracts)

Palmquist, R. A. and Kim, K. S. (2000). Cognitive style and on-line database search experience as predictors of web search performance. *Journal of the American Society for Information Science*, 51, pages 558-566.

Wang, P., Hawk, W. B., and Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management*, 36, pages 229-251.

Wolfram, D. and Dimitroff, A. (1998). Hypertext vs. boolean-based searching in a bibliographic database environment: A direct comparison of searcher performance. *Information Processing and Management*, 34, pages 669-679.

Interpersonal Attraction from Electronic Self-Disclosure: Cumulative Effect of CMC

Hiroaki Morio
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113-0033
Japan
hmorio@l.u-tokyo.ac.jp

Deborah Richardson
Augusta State University
2500 Walton Way, Augusta
Georgia 30904
USA
drichardson@aug.edu

Bibb Latanè
Center for Human Science
212 Vance Street, Chapel Hill
North Carolina 27516
USA
latane@humanscience.org

Abstract

As a part of a weekly computer lab, students from three courses exchanged messages on personal and political topics with neighbors in an experimentally controlled cyberspace for seven weeks to investigate the effect of message exchange on development of interpersonal attraction. Personal topics were drawn from an experiment on interpersonal closeness (Aron et al., 1997); political topics were drawn from Machiavellianism scale (Christie and Geis, 1970). After message exchange, participants perceived less similarity with their neighbors on personal than political topics, perhaps because personal topics provided more individuating information. This difference in perceived similarity was not associated with ratings of interpersonal attraction. Number of messages exchanged did not have a large effect on the weekly ratings, but related strongly to the final ratings of neighbors. The results suggest the importance of longitudinal studies of interpersonal attraction.

Keywords: interpersonal attraction, CSCL, CMC, self-disclosure

1 Introduction

1.1 Interpersonal relationship and Social Intelligence Design

When we think of Computer-mediated communication (CMC) media, such as e-mails and electronic forums, as tools for building and maintaining interpersonal relationships, one obvious feature is that most of nonverbal information available in Face-to-Face (FtF) interaction, such as facial expressions and tone of voice, is not available via typical text-based CMC. Researchers have argued that this lack of nonverbal cues result in reduced social presence in CMC (Sproull and Kiesler, 1991; Kiesler et al., 1984). It is also called as "cues-filtered-out" effect (Culnan and Markus, 1987), because CMC can be considered to "filter out" any nonverbal cues.

Early research indicated that this lack of social-cues make CMC characteristically less personal, less emotional, and unsociable compared to FtF communication (Hiltz et al., 1986). On the other hand, Walther and his colleagues (1992, 1995; Tidwell and Walther, 2002) have repeatedly shown that the effect of lack of social-cues in CMC on interpersonal relationship depends on the context. They criticized that the earlier studies demonstrating CMC as impersonal medium exclusively

used experimental settings with one-shot interactions among strangers, and provided some evidence that CMC can be personal and socioemotional (see Walther, 1995).
Building and maintaining interpersonal relationship is itself important function of communication, but also plays a key role in task-oriented groups or communities. Numerous studies on group dynamics has emphasized the importance of interpersonal relationship in groups, and it is also true in online groups. For example, Kreijins et al. (2003) argued that relationship building processes among participants are essential part of an effective Computer-Supported Collaborative Learning (CSCL) system. As Social Intelligence (SI) is defined as the ability for people to relate to, understand and interact effectively with others (Fujihara, 2001), understanding how people build relationships would lead to designing a more effective system. In this study, we explore possible factors that would influence interpersonal relationship built via CMC, and discuss the implications of such factors in building a better system.

1.2 Cumulative effect of CMC in interpersonal relationship

Unlike FtF communication, CMC does not have to be synchronous. It can be asynchronous: With e-mails and electronic "bulletin board", users can send messages to those who are not present at the moment. At the same time, the technology allows users to correspond to each other almost immediately if the communication partner is available. If we examine the relationship-building function of CMC, we have to take these features into consideration, as well as "cues-filtered-out" approach mentioned before. Walther (1992) proposes that the chronological development of interpersonal relationships via CMC might be different from FtF relationships. He argues that it takes longer for people to process socioemotional information from verbal and textual messages in CMC than from multichannel messages in FtF; therefore, it takes longer to develop a relationship. Field studies have been supporting this proposition (see Walther, 1992), yet most experimental studies that have been utilizing one-shot design have been failing to confirm it.

1.3 Knowledge Networking for Teaching in Science

The goal of this study is to examine a long-term effect of CMC on formation of interpersonal relationships in a controlled setting. We used a CSCL system with a specific message control scheme called KNITS (Knowledge Networking for Teaching in Science). KNITS was built based on earlier experiments on electronic communication, in which university students exchanged e-mail messages about a variety of topics within controlled communication structures (Latanè and Bourjeois, 1996; Latanè and L'Herrou, 1996). With KNITS, people are distributed in a hypothetical two-dimensional electronic village. Figure 1 shows the geometry used in this study called Ribbon Geometry (Latanè, 1996). Because the amount of, and the nature of, communication is determined by the distance between people in this village, it is possible to manipulate various aspects of interpersonal interaction. On the other hand, interaction in KNITS is neither reciprocal nor voluntary. The flow of messages is systematically controlled by a pre-arranged scheme based on distance among participants. These characteristics make social interactions in KNITS highly artificial compared to CMC in real life settings. We can, however, manipulate factors to see their effects without confounding with other factors.
The first factor we examine in this study, is an effect of the amount of communication on interpersonal attraction. A well-established theory in social psychology called "mere exposure effect" suggests that more frequent exposure to an object leads more favorable ratings of the stimuli (Zajonc, 1968). Applying this theory to KNITS paradigm, we expect that more interaction leads to more attraction.

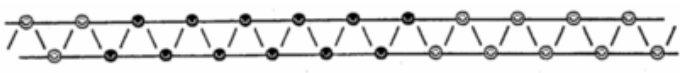


Figure 1: An Example of Virtual Electronic Village: Ribbon Geometry

1.4 Electronic Disclosure Game

In literature dealing with FtF interpersonal relationship , self-disclosure is considered as an essential part of relation-building process (see Brehm, 1992). Because text-based CMC does not convey nonverbal social cues, people would need to play more attention to send and receive personal information in CMC than in FtF to build relationship (Joinson, 2001; Tidwell and Walther, 2002). To control self-disclosure of participants, a procedure devised by Aron et al. (1997) was adopted for KNITS in this study. Aron et al. (1997) proposed an experimental methodology called "disclosure game" to study people's interpersonal relationship. Their experimental procedure allows participants to conduct self-disclosure on various pre-arranged topics, designed to increase their intimacy. Our electronic version of "disclosure game" also asked participants to discuss specific topics given in order designed to increase intimacy.

In addition to these personal topics designed to promote intimacy, we also prepared political/philosophical topics taken from Machiavellianism scale (Christie and Geis, 1970). Machiavellianism scale is a psychological test to measure how much people agree with the principles of Niccolo Machiavelli that any means may be justifiably taken in order to achieve one's goal. By mixing and matching personal topics and political topics, we could examine the effect of content of communication, personal vs. political disclosure, as well as the effect of homogeneity of the content.

2 Methods

2.1 Participants

As a part of weekly computer laboratory, 135 undergraduates (113 females and 22 males) from three social psychology courses participated in the experiment. In the first week, students were randomly assigned to a node in a one-dimensional social space, or an electronic village.

2.2 Procedure

On Week 1 through 7, students wrote messages on discussion topics chosen to permit personal or political/philosophical disclosure. Twenty personal topics were taken from Aron et al. (1997) (e.g., "Given the choice of anyone in the world, who would you want as a dinner guest, and why?") Topics were taken only from the first set, which were designed to introduce participants into interpersonal disclosure. Twenty political topics were taken from Machiavellianism scale by Christie and Geis (1970) (e.g., "Never tell anyone the real reason you did something unless it is useful to do so.") These messages were written as a response to a question item. Therefore, participants produced only one message per item and, the messages were not addressed to specific individuals.

Topic rating task was added to see preference of participants for message exchange subject. On Week 6, students rated eight randomly selected topics out of 40 for importance, intimacy, and how much they would like to read and to write about them. These 40 topics included 20 personal topics and 20 political topics. Eighteen topics were used for the message exchange while the rest of 22 topics were not used for the message exchange.

On Week 2 through 7, the messages written in the previous week were delivered to those who are "neighbors" in the electronic village according to the message exchange plan. Neighbors were identified only by 3-character aliases throughout the experiment. A hypothetical example of how messages would be delivered to their neighbors from week n to week n+1 is shown in Figure 2. The letters A through F donates individuals. In this example, students read messages written in a previous week by other students with relative addresses of -2, -1, and +1.

Table 1 shows details of when and to whom messages were delivered each week. A cell with a combination of a number and an alphabet represents a delivery of message from a neighbor to a participant in a week. The number in each sell represents a number of messages delivered

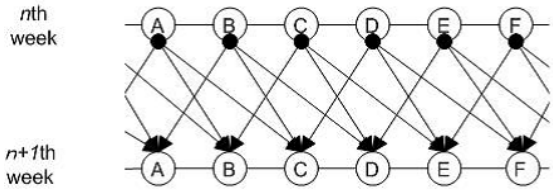


Figure 2: An Example of Message Flows in a Week

from the neighbor, while the alphabet shows the type of a topic(s). For example, in week 4, a participant received one message on political topics from a neighbor of relative address -6, four messages on political topics from a neighbor of relative address -2, four messages on personal topics from a neighbor of relative address +2, and one message from a neighbor of relative address +6.

Table 1: Numbers and Types of Message Exchange by Relative Address and Week

	relative address											
Week	*-6*	*-5*	*-4*	*-3*	*-2*	*-1*	*1*	*2*	*3*	*4*	*5*	*6*
1												
2			**2c**			**4B**			**2B**	**2a**		
3		**1c**		**2B**			**8B**				**1a**	
4	**1c**				**4c**			**4a**				**1a**
5			**2c**	**2a**		**4B**			**2B**			
6		**1c**		**2B**			**8B**				**1a**	
7				**2c**	**4c**			**4a**		**2a**		
total	**1c**	**2c**	**4c**	**8B**	**8c**	**8B**	**16B**	**8a**	**4B**	**4a**	**2a**	**1a**

The message exchange was planned so that nearer neighbors send and receive more messages. In addition, the messages from the neighbors on the right side (with positive relative addresses) are on personal topics, while those from the left side (with negative relative addresses) are on political topics. Students could receive 1, 2, 4, or 8 messages from a neighbor on either personal, political, or both personal and political topics each week. As can be seen from Table 1, The message exchange was not necessarily reciprocal.

Immediately after reading each message set, participants rated that neighbor for perceived similarity (1. not at all similar to me ... 4. very similar to me) and liking (1. don't much care for ... 4. like a lot) on a 4-point Likert scale each week. In addition to these weekly ratings of neighbors, students were asked to rate all of their 12 nearest neighbors that they communicated with on familiarity (1. don't know at all ... 4. feel I know well) and liking (1. don't much care for ... 4. like a lot) at the end of all the message exchange (Week 7). These weekly ratings and final ratings of neighbors were used as indicators of interpersonal attraction developed among participants.

3 Results

3.1 Overview of Analysis

First, students' ratings of topics will be examined to see which kind of topics students prefer for the message exchange in the context of electronic interaction with classmates. It is predicted that personal topics are preferred to political topics. Next, Weekly and final assessments of interpersonal attraction will be examined. The design of this study enables us to investigate the

development of personal attraction each week as well as the interpersonal attraction formed as a result of communication for the length of seven weeks. There are three factors systematically manipulated in the message exchange scheme. These three factors lead to the following three working hypotheses we can test by creating proper contrasts.

1. Topic type (Personal topics produces more interpersonal attraction than political topics.)
2. Heterogeneity of topics (Messages based on both personal and political topics lead to more interpersonal attraction compared to either personal topics only or political topics only.)
3. Number of messages (The more messages one receives from a neighbor, the more interpersonal attraction.)

These tests can be performed on the weekly ratings as well as the final ratings of interpersonal attraction. As these effects are cumulative, it is predicted that the effects are stronger for the final ratings than the weekly ratings.

3.2 Manipulation check: Topic ratings

As each student rated randomly selected 8 out of 20 topics, number of participants rated each item ranges from 18 to 42. Mixed-design Analysis of Variance (two between factors, topic type and whether used in discussion or not) was conducted, with item as a nested factor under topic type on student's ratings on topics. Personal topics were rated as more important, $F(1, 1024) = 10.55, p < .001$, more intimate, $F(1, 1024) = 57.17, p < .001$, more like to read, $F(1, 952) = 15.91, p < .001$, and more like to write, $F(1, 952) = 23.63, p < .001$ than political topics. Table 2 shows means and standard deviation of four ratings by topic types.

Table 2: Means and standard deviations for topic ratings by topic types

Topics		*Importance*	*Intimacy*	*Liking to read*	*Liking to write*
Personal topics	*M*	3.25	2.90	3.11	2.91
	SD	.96	1.12	1.11	1.15
Political topics	*M*	3.07	2.43	2.84	2.58
	SD	.96	1.05	1.10	1.13

3.3 Weekly neighbor ratings

For the effects of topic type, heterogeneity of topics (whether messages are only on personal or political topics or on both types of topics), and number of messages received from the target of ratings in the week, effect sizes were calculated for each week. The effect sizes, in a form of zero-ordered Pearson product-moment correlation coefficient r, were obtain by calculating d indices from the means and the standard deviations. Neighbors who wrote on political topics were consistently perceived as somewhat more similar although not more likable (average r across weeks = .14 *vs.* .02), than neighbors who wrote only on personal topics. Topic heterogeneity did not influence perceived similarity or liking. The number of messages in a given week had little effect on immediate reactions. It was significant only in week 2, even though its size was small ($r = .21, p < .05$, for liking, and $r = .20, p < .05$, for similarity). In the following weeks, the effect of number of messages was not significant on either of the ratings ($rs < .15$, *ns.*) In general, the three factors had a little or no impact on weekly ratings of interpersonal attraction.

3.4 Final assessment of interpersonal attraction

Repeated measure one-way ANOVA showed a significant effect of relative addresses on both perceived similarity, $F(11, 1463) = 15.99, p < .001$, and liking, $F(11, 1463) = 11.95, p < .001$. These

omnibus F-test results show that interpersonal attraction toward neighbors formed at the end of the last week was significantly influenced by their relative addresses. In the following sections, each of three factors described in the overview section is individually examined with a particular contrast devised to test the hypothesis.

3.4.1 Topic Type

A contrast was devised to compare interpersonal attraction toward all the neighbors who sent messages on personal topics (relative addresses: -6, -5, -4, and -2) and those who sent messages on political topics (relative addresses: +6, +5, +4, and +2). Contrary to what was found in the results of weekly ratings, two separate t-tests showed that the effects of topic type on the either perceived familiarity, $F(1, 133) = 1.26, p > .26$, or liking, $F(1, 133) = 2.551, p > .11$, were not significant after completing all the message exchanges (see Table 3 for means and standard deviations.) Communicating on either personal or political topics did not make any difference in the formation of interpersonal attraction.

Table 3: Effect of topic type on the final assessment of interpersonal attraction

	Topic type	
Ratings	*Personal*	*Political*
Familiarity		
M	1.96	1.90
SD	.63	.65
Liking		
M	2.35	2.29
SD	.56	.57

3.4.2 Heterogeneity of topics

Another contrast was devised to see the effect of heterogeneity of topics. The contrast compares neighbors who sent messages on heterogeneous topics (relative addresses: -3 and -1) with those who sent messages on homogeneous topics (relative addresses: -2 and 2). Only the neighbors sending eight messages were selected for this contrast to avoid confounding the effect of heterogeneity with the effect of number of messages. The result of the t-test was significant for neither perceived familiarity, $F(1, 133) = .19, ns$, nor liking $F(1, 133) = 1.468, ns$. Heterogeneity of topics did not have influence on interpersonal attraction either.

3.4.3 Number of messages

During the course of seven weeks, a neighbor could send either one, two, four, eight, or sixteen messages to a participant depending on his/her relative address to the recipient. As the comparison is not of pairwise, composite scores were created for each of total number of messages. Repeated measures ANOVA (total number of messages as within-subject factor) was then conducted to test the effect. The effect was significant on both perceived familiarity, $F(4, 532) = 41.22, p < .001$, and liking, $F(4, 532) = 22.08, p < .001$.

Figure 3 illustrates that higher ratings on perceived familiarity and liking were obtained for those who sent more messages. Not only that the result supported DSIT's prediction that greater exposure led to more attraction, the relationship between total number of messages and interpersonal attraction fit Latanè's Social Impact Theory's Psychosocial Law: social influence increases in proportion to some root of the number of sources (Latanè, 1981).

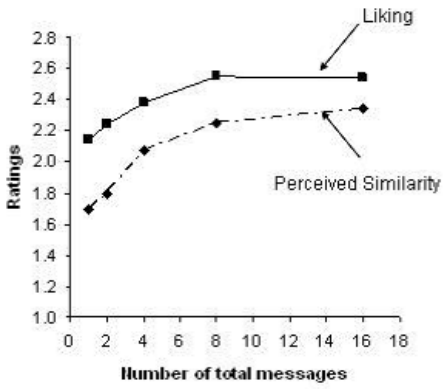


Figure 3: The Effect of Number of Total Messages on Interpersonal Attraction

4 Discussion

The results clearly demonstrated that the interpersonal attraction formed in a relatively longer span, such as several weeks, cannot be understood by observing single instances of interactions. In this study, number of messages from a person had a little, if any, effect on weekly interpersonal attraction ratings, yet total number of messages exchanged was the primary determinant of final interpersonal attraction. It is also worth noting that small differences in weekly similarity ratings due to topic type did not have any effects on final liking ratings. In sum, interpersonal relationship developed through a series of interaction was not equal to simple aggregation of individual pieces of interactions. These findings support the proposition that development of interpersonal relationship via CMC should be studied chronologically over a relatively long period of time.

Another lesson learned here is the importance of the amount of communication, not the content. The two types of topics provided in this study were perceived differently by students: Students liked to read and write about personal more than political items and saw them as more important and intimate. In addition, topic type had influence on impression formation on a weekly basis. Participants perceived less similarity with their neighbors on personal than political topics, perhaps because personal topics provided more individuating information. However, topic type had little effects on both weekly ratings and final attraction ratings. It contradicts with findings in the literature of FtF interpersonal attraction: one would expect messages involving self-disclosure would better serve as a tool for building a stronger relationship. Other studies in CMC suggests that people do engage in various tactics of self-disclosure online (Tidwell and Walther, 2002). As self-disclosure of participants in this study was rather forced as responses to topics, and not spontaneous messages aimed to someone in particular, we do not deny the possibility that spontaneous self-disclosure in CMC has a positive effect. If perception of the intent of message is what counts most, self-disclosure in this study is not qualified for this purpose. In this regard, more studies are needed to elucidate people's spontaneous self-disclosure and its effect on interpersonal relationship.

In KNITS, sheer amounts of communication controlled by the system predicted the development of interpersonal relationships among participants, and the content of the communication did not. Regardless of the content, the more students communicated, the more they liked each other. This partially explains how a web of strong interpersonal relationships, or a community, can develop in an anonymous yet identifiable CMC environment, such as electronic bulletin boards or mailing lists. In these environments, people usually know about each other very little, and the communica-

tion rarely contains any self-referencing information, or "wet information", following classification scheme of Information Humidity Theory (Azechi, 2000). The result of this study suggests that exchange of dry information itself is sufficient for people to develop interpersonal attraction for each other.

The interpersonal interaction occurred in this experiment was quite artificial, and the results are not to be blindly applied to real-life situations. However, because the interactions were highly regulated, the system in this study enabled us to examine whether the content and the amount of communication have impact on interpersonal attraction independently. In real-life situations, these two factors are interdependent and intertwined: the amount of communication would influence its content, and vice versa. As we have entangled their effects on interpersonal attraction in CMC to some degree here, the next step to be taken is to examine how they influence each other, in a more naturalistic environment.

Insight gained from this study regarding Social Intelligence Design is summarized as this: simply increasing the amount of information is the first step for improving interpersonal attraction among users. Increased interpersonal attraction will leads to a more cohesive group and eventually a community effectively serving its goal.

References

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., and Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23:363–377.

Azechi, S. (2000). Social psychological approach to knowledge-creating community. In Nishida, T., editor, *Dynamic Knowledge Interaction*, pages 15–57. CRC Press.

Brehm, S. S. (1992). *Intimate relationships (2nd ed.)*. New York: Mcgraw-Hill.

Christie, R. and Geis, F. (1970). *Studies in Machiavellianism*. New York: Academic.

Culnan, M. J. and Markus, M. L. (1987). Information technologies. In Jablin, F. M., Putnam, L. L., Roberts, K. H., and Porter, L. W., editors, *Handbook of organizational communication: An interdisciplinary perspective*, pages 420–443. Newbury Park, CA: Sage.

Fujihara, N. (2001). How to evaluate social intelligence design. In Terano, T., Nishida, T., Namatame, A., Tsumoto, S., Ohsawa, Y., and Washio, T., editors, *New Frontiers in Artificial Intelligence - Joint JSAI 2001 Workshop Post-Proceedings, Lecture Notes in Artificial Intelligence*, page 2253. Springer-Verlag.

Hiltz, S. R., Johnson, K., and Turoff, M. (1986). Experiments in group decision making: Communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research*, 13:225–252.

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, 331:177–192.

Kiesler, S., Siegel, J., and McGuire, T. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39:1123–1134.

Kreijins, K., Kirschner, P. A., and Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, 19:335–353.

Latanè, B. (1981). The psychology of social impact. *American Psychologist*, 36:343–365.

Latanè, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication*, 46:13–25.

Latanè, B. and Bourjeois, M. J. (1996). Experimental evidence for dynamic social impact: The emergence of subcultures in electronic groups. *Journal of Communication*, 46:35–47.

Latanè, B. and L'Herrou, T. (1996). Social clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of Personality and Social Psychology*, 70:1218–1230.

Sproull, L. and Kiesler, S. (1991). *Connections: New ways of working in the networked organization*. Cambridge: MIT Press.

Tidwell, L. C. and Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, 28:317–348.

Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19:52–90.

Walther, J. B. (1995). Relational aspects of computer-mediated communication: experimental observations over time. *Organizational Science*, 6:186–203.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9:1–27.

Discovery of Web User Communities from Client-level Log Data *

Tsuyoshi Murata
National Institute of Informatics & Japan Science and Technology Agency
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan
tmurata@nii.ac.jp

Abstract

There are two kinds of communities in the Web; communities of related Web pages (Web communities) and communities of users who watch such related pages (user communities). Discovery of the former communities has been attempted by many researchers such as Kumar's trawling and Flake's method. Discovery of the latter communities is also important for clarifying the behaviors of Web users. Moreover, it is expected that the characteristics of user communities in the Web correspond to those in real human communities. A method for discovering user communities is described in this paper. Client-level log data (Web audience measurement data) is used as the data of users' Web watching behaviors. Maximal complete bipartite graphs are searched from the graph obtained from the log data without analyzing the contents of Web pages. Experimental results show that our method succeeds in discovering many interesting user communities with labels that characterize the communities.

Keywords: Web usage mining, Discovery, User Communities, Web Audience Measurement Data

1 Introduction

World Wide Web is an important media that allows us to distribute messages with Web pages, and to refer others by hyperlinks. The Web has abilities of uniting related Web pages as well as humans of similar tastes. The former (groups of related Web pages) is often called Web communities, and the latter (groups of users of similar tastes) is called user communities in this paper. Web communities are based on the connection of pages with hyperlinks, and user communities are based on the users' behaviors of watching Web pages. Both communities are mutually related: 1) if many users have interests to some specific topic, the number of Web pages about the topic is increased, and 2) if the structure of Web pages about specific topic has been changed, users' Web watching behaviors will be affected. Discovering the structure of both communities and the interactions between them is important for predicting the development of the Web. Several methods have been proposed for the discovery of Web communities, such as Kumar's trawling based on graph search(Kumar, 1999) and Flake's method based on network-flow theory(Flake, 2002). Compared with the research for Web communities, however, little attention has been given to the research for user communities in the Web. There are many practical applications for discovering user communities, such as recommendation of suitable Web pages and adaptation of Web sites for users of similar tastes.

This paper proposes a method for discovering user communities from the data of users' Web watching behaviors. Discovery of user communities are important for understanding users' information needs, and for recommending suitable Web pages by social Web filtering. Web audience measurement data are client-level log data that record users and their visited URLs. The data are

*The author would like to express his thanks to the reviewers for giving fruitful comments, and to Prof. Takeaki Uno for permitting the use of his fast bipartite search algorithm.

time	userID	elapsed time	URL
00:00	9601	10	www.jpncm.com/cgi-lib/cmbbs/wforrum.cgi
00:00	9701	27	www.dion.ne.jp
00:00	3502	19	search.auctions.yahoo.co.jp/search
00:00	5201	14	eee.eplus.co.jp/shock/shock03.html
00:01	5502	10	user.auctions.yahoo.co.jp/jp/show/mystatus
00:01	0501	6	user.auctions.yahoo.co.jp/show/mystatus
00:01	3301	36	www.pimp-sex.com/amateur/raimi/01/clean.htm
00:01	9701	4	auctions.yahoo.co.jp/jp/2...-leaf.html
00:02	8501	3	www.uicupid.org/chat/csp-room.php
00:02	8001	3	page.auctions.yahoo.co.jp/jp/show/qanda
00:02	1501	11	www.nn.iij4u.or.jp/ movie/pm/main.html
00:02	9002	12	www.umai-mon.com/user/p-category.php

Figure 1: Example of Web Audience Measurement Data

transformed into a graph, and complete bipartite graphs that correspond to user communities are searched from the graph. Although the method is quite simple, our system based on the method succeeds in discovering many interesting user communities.

2 Web Audience Measurement Data

Web audience measurement data are just the same as audience data of TV programs. Randomly selected users are asked to use special modified Web browsers that record users' behaviors (such as visited URL, time of the visit, and elapsed time at the URL) at users' PC. Example of Web audience measurement data are shown in Fig. 1.

Each row of Web audience measurement data represents a user's visit of a URL. Attributes of the data include time, user ID, visited URL, elapsed time, and other miscellaneous information about users' actions such as click of hyperlinks, back to previous pages, or manual enter of URLs.

In Japan, there are four major companies for this sort of investigation: Nielsen//NetRatings (`http://www.netratings.co.jp/`), VideoResearch Netcom (`http://www.vrnetcom.co.jp/`), Nikkei BP (`http://www.nikkeibp.co.jp/`), and Nihon Research Center (`http://www.nrc.co.jp/`). (Experimental results of this paper are based on the data from one of these companies.) Web audience measurement data are used mainly for statistical analysis, and for detailed investigation of the visitor of specific site. For example, 1) Investigation of the situations of internet usage (users' age or gender, access time, environment of users' computers, and so on), 2) Investigation of the relation between campaigns of sales promotion of a company and the number of visitors to the company's Web site, and 3) Investigation of the relation between behaviors of buyers at online shops and the results of their questionnaire. Fig. 2 shows the example of users' personal data. The data contain attributes such as user ID, gender, age, birth year, birth month, occupation, address, and so on.

In general, data source for Web usage mining can be divided into two classes: server-level data and client-level data(Srivastava, 2000). Web audience measurement data is the latter. As is often pointed out, usage data at client-level reflect users' true behaviors since the usage of cached data cannot be recorded on server-level data.

3 A Method for Discovering Web User Communities

As the method for discovering user communities, graph mining approach is employed in this paper. This is because of the following reasons: 1) in order to analyze huge Web usage data, graph mining approach is generally faster than content based approaches, and 2) it is expected that methods for Web structure mining are applicable for graph representation of Web usage data.

userID	gender	birth year	birth month	job	area
0016	M	1971	9	22	3
0017	M	1981	9	74	3
0019	M	1939	12	94	3
0020	M	1950	11	21	3
0021	F	1980	3	75	3
0022	F	1976	12	95	3
0023	F	1975	7	96	3
0024	M	1945	5	41	3
0025	M	1963	12	13	5
0026	M	1960	11	41	3
0027	M	1971	4	11	3
0028	F	1946	8	81	3
0029	M	1944	9	42	3
0030	M	1975	9	75	3
0031	F	1976	4	82	3

Figure 2: Example of Users' Personal Data

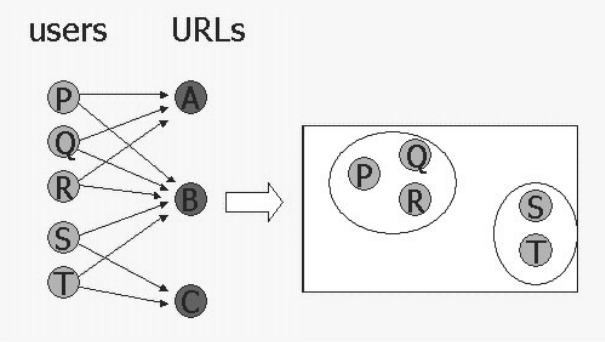


Figure 3: Discovery of User Communities from User-URL Graph Structure

Several approaches have been attempted for the discovery of Web communities, which is one of the important research topics of Web structure mining. Kumar (Kumar, 1999) claimed that Web pages whose hyperlinks constitute a bipartite graph structure are regarded as a community sharing common interests. The idea is simple and powerful for finding related nodes from a graph, and it is expected to be applicable to the graph structure of Web usage data. The initial idea of our approach is to search complete bipartite graphs from the pairs of user and visited URL. Let us suppose group of users (P, Q, R, ...) visit URLs (A, B, C, ...), as shown in Fig. 3. Users' Web watching behaviors can be represented as a graph when we regard each user and URL as a node and each visit of URL as an edge. Bipartite graphs such as (P, Q, R, A, B) and (S, T, B, C) can be regarded as user communities sharing common interests since the users in the graph visit the same URLs.

However, this naive idea is not applicable to real Web audience measurement data since most of the URLs are visited by only one user. The graph structure of users and URLs (user-URL graph) is sparse and most of the complete bipartite graphs contained in the graph are composed of only few nodes. As a method for discovering user communities from the structure of Web audience measurement data, the following procedures are applied in this paper:

1. Decomposition of URLs into terms

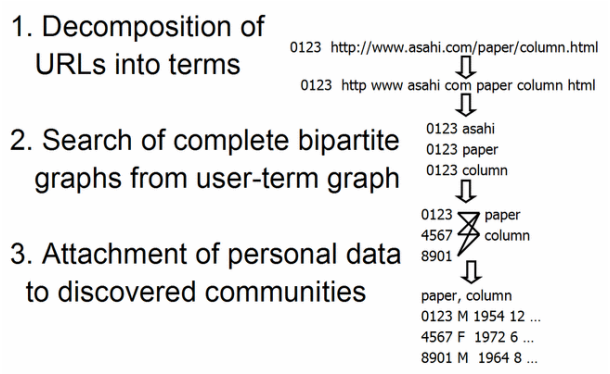


Figure 4: A Method for Discovering User Communities

A URL contains information about the page it points to. Each URL string is decomposed into terms that are used as labels of the URL.

2. Search of maximal complete bipartite graphs from user-term graph
   As claimed by Kumar, Web pages whose hyperlinks compose a complete bipartite graph are mutually related. The same idea is applied to the user-term graph.

3. Attachment of personal data to discovered user communities
   In order to assist the understanding of discovered user communities, personal data (such as gender and birth year) of all the members are attached to user communities.

Overall procedures are shown in Fig. 4.

3.1 Decomposition of URLs into terms

As mentioned above, each URL contains useful information about the page it points to. For example, URL of SID2004 conference is `http://parlevink.cs.utwente.nl/sid04.html`, and terms included in the URL string (such as utwente and sid04) can be used as the labels for the Web page of the URL. Another reason for using URL strings for discovery is that many Web pages are dynamically generated at the time of visit. Contents of such pages cannot be obtained at the time of discovery because of privacy reasons and storage capacity reasons. It is expedient to use URL strings as the labels of huge Web log data. Decomposition of URLs is performed as follows:

1. Terms are extracted from URL strings by chopping URL at the point of punctuations, such as periods (.), slashes (/), dashes (-), and question marks (?).

2. Overly frequent terms (such as “www” and “html”) and numbers are removed from the extracted terms because they do not characterize the contents of original URLs. In our experiment, the following terms are removed: “www”, “html”, “htm”, “cgi-bin”, “cgi”, “co”, “ne”, “or”, “jp”, “net”, and “com”.

3. By replacing URLs with the terms generated in the above procedure, a graph of userIDs and URLs (user-URL graph) are expanded to the graph of userIDs and terms (user-term graph). In general, obtained user-term graph is denser than original user-URL graph since more than one term are usually extracted from each URL.

3.2 Search of Maximal Complete Bipartite Graphs

Generated user-term graph is a fairly large bipartite graph. From the graph, maximal complete bipartite graphs are searched. Such graphs indicate that a group of users visit URLs that include a group of terms. Search of bipartite graph is performed by the algorithm proposed by Uno(Uno,

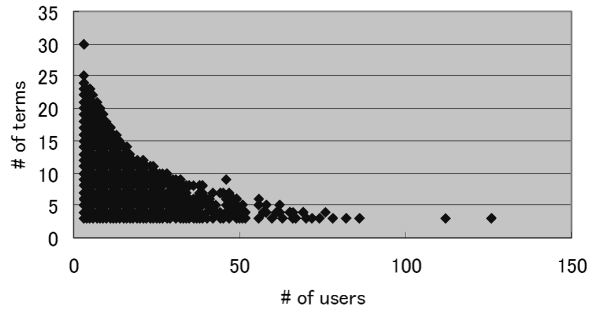


Figure 5: Distribution of the Size of User Communities

2003). The algorithm has abilities of searching each complete bipartite graph with the time complexity O(D^3), where D represent the maximum number of degree in a graph. For the search of user-term graph, the algorithm is much faster than conventional search methods whose time complexity are O($|V||E|$) at best. In our experiment, maximal complete bipartite graphs which are equal or larger than (3, 3) graph are searched since there are many small isolated complete bipartite graphs (such as one-to-one connection of one user and one term).

3.3 Attachment of Personal Data to UserID

Many complete bipartite graphs are obtained by the above search. Each complete bipartite graph is shown as a list of userID and terms. As a way to assist the understanding of a user community, personal data that correspond to each userID are attached to resultant graphs. Therefore, each user community is finally represented as terms and personal data (such as gender and birth year) as shown in Fig. 2.

4 Experiments

Based on the method described above, experiment for discovering user communities is performed. Web audience measurement data used in our experiment are obtained from 6046 users during January 24-28, 2001.

Fig. 5 shows the distribution of the size of discovered user communities. You can see from the figure that the more users included in a user community, the less the number of shared terms is. Most of the discovered user communities have less than 50 users. Reasonable size of user communities is controversial and it depends on how the communities are used. At least, we can claim that the method has abilities of discovering several sizes of unobvious user communities.

Evaluation of the quality of user communities is not an easy task. Communities that contain meaningful terms are selected by manual inspection (mere meaningless strings are often included as terms). Some of the interesting results are shown below. As explained above, each user community is represented as a collection of terms and personal data of users (gender, birth year, occupation, and address).

User Community about Soccer

Terms:

- “news”
- “soccer”
- “nikkansports” (name of Japanese sports newspaper company)

Users:

- Male, 1980, college student, kanto
- Male, 1976, college student, kanto
- Male, 1965, service job, kanto
- Male, 1972, college student, kanto

The above terms and users mean that the four users visited URLs that contain the terms ("news", "soccer", and "nikkansports") as their substrings. Since such URLs are expected to point to Web pages about soccer, these users are regard as a community sharing common interests about soccer. These users are young men in their twenties and thirties judging from their birth year. It seems reasonable that these users are interested in "soccer" and "news".

User Community about Money

Terms:

- "news"
- "yahoo"
- "search"
- "money"
- "nikkei" (name of Japanese financial newspaper company)

Users:

- Male, 1957, transportation, chubu
- Male, 1941, none, tokai
- Male, 1956, teacher, kyushu
- Male, 1944, none, kanto

Since this user community contains "money" as one of its terms, this is regarded as a user community regarding money. As you can see, the uses in this community are elder men (from forties to sixties) and two of them have no job (probably already retired). These members are completely different from those of the above "soccer" community. Our method succeeds in discovering various kinds of user communities by graph search, and the communities are labeled with their terms. This makes the understanding of user communities easier than simple user clustering.

There are many other interesting user communities discovered with our method. Brief explanations about some of them are described below:

- The biggest user community (126 users) has the following terms: "index", "yahoo", and "search". Probably these users visit Yahoo and perform search. It is controversial to regard such users as a community. However, it is reasonable that the number of users of this community is the biggest.
- A user community about university entrance exam and preparatory school is discovered. All the members of the community are around 20 years old, which are valid in Japanese educational system.

5 Discussion

5.1 Related work of user community

There are several attempts for the research of Web log mining such as discovering frequent patterns of log data, or clustering users of specific Web site. Our approach is quite different in that the goal is to discover user communities of similar tastes using Web audience measurement data that record users' visits of several Web servers.

As the research for discovering humans' social network, ReferralWeb (Kautz, 1997) is a famous one. Co-citation of technical papers is used for finding relations among AI researchers. Although Kautz's approach focuses on the network among humans, our method utilizes both users and terms, which enables the discovery of user communities that are labeled with terms.

Another related approach is social network analysis from graph structure (Girvan, 2001)(Tyler, 2003). In this approach, given graphs are decomposed into subgraphs by removing unnecessary edges based on a criterion called edge betweenness. Although the approach succeeds in discovering fractions in a sports club and informal groups in a company, it is not applicable to huge Web data because of its scalability.

In general, discovery of user communities are important for fostering communication among users (such as I2I project (Budzik, 2002)), and for performing collaborative filtering (such as Grouplens project). User communities discovered by our new method is rather unsubstantial; each member does not know which communities he or she belongs. Such user communities may not be appropriate for fostering explicit communication among users; they can be used for implicit collaborative filtering (for users) and for site modification (for Web site administrators).

5.2 Evaluation of User Communities

One of the most crucial problems for the research of community discovery is to evaluate its outputs. As mentioned above, evaluation of user communities is not an easy task. In Girvan's paper(Girvan, 2001), several examples of communities are shown. Tyler(Tyler, 2003) attempts interviews to human subjects in order to evaluate discovered communities.

Evaluation of user communities depends on the way they are used. Our ultimate goal for user community discovery is to clarify dynamic changes of user communities. As well as real human communities, user communities in the Web are expected to born, grow, and decay over time. In this paper, distribution of the size and the number of user communities are shown in Fig. 5, and some examples of user communities are shown in section 4. Personal data such as gender or birth year are attached to each user community, which make the understanding of user communities easier. Moreover, since our method finds bipartite graph of userID and terms, the terms can be regarded as labels of the communities. Such labels are useful for tracing dynamic changes of user communities over time.

5.3 Web Communities and User Communities

Discovery of user community proposed in this paper is the first step for clarifying dynamic changes of user communities. It should be mentioned that Web communities and user communities are closely related, and that some changes of one side may affect the other. Clarifying the interactions between both communities is important for retrieving information from the Web as well as for supporting human relationship through the Web.

As the bridge for connecting both communities, search engines often play an important role. For example, when a worldwide accident such as the 9.11 terror occurred, many search of related words such as "CNN" or "world trade center" were performed immediately after. As the next step, information exchange on bulletin boards becomes active, and Web pages about the accident are newly built and linked together. Detailed analysis of the process of this kind of community generation is expected to assist humans' smooth communications through the Web.

6 Conclusion

This paper describes a method for discovering user communities from Web audience measurement based on the search of maximal complete bipartite graphs. Many interesting user communities are actually discovered. Graph mining approach is attempted mainly for Web structure mining. One of the contributions of this paper is that the approach is effective also for Web usage mining. There are many interesting and challenging research topics as the next step of this research, such as detecting relations among different user communities, and analyzing dynamic changes of user communities.

References

Budzik, J., Bradshaw, S., Fu, X., and Hammond, K. J., (2002). Clustering for Opportunistic Communication, in Proc. of WWW 2002.

Flake, G. W., Lawrence, S., Giles, C. L., Coetzee, F. M. (2002). Self-Organization and Identification of Web Communities, IEEE Computer, Vol. 35, No. 3, pages 66–71.

Girvan, M., Newman, M. E. J. (2001). Community structure in social and biological networks, online manuscript, `http://arxiv.org/abs/cond-mat/0112110/`.

Kautz, H., Selman, B., Shah, M. (1997). The Hidden Web, AI Magazine, Vol.18, No.2, pages 27–36.

Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A. (1999). Trawling the Web for Emerging Cyber-Communities, Proc. of the 8th WWW conference.

Srivastava, J., Cooley, R., Deshpande, M., Tan, P.-N. (2000). Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, ACM SIGKDD Explorations, Vol.1, No.2, pages 12–23.

Tyler, J. R., Wilkinson, D. M., Huberman, B. A. (2003). Email as spectroscopy: automated discovery of community structure within organizations, `http://xxx.lanl.gov/arXiv:cond-mat/0303264`.

Uno, T. (2003). Fast Algorithms for Enumerating Cliques in Huge Graphs, submitted.

Community Site Collaboration Type Broadcasting System Using Interface Agent and Multimedia

Toshihiro Murayama
Research Institute of Science and Technology for Society, JST
ATAGO green hills MORI Tower 18F, Atago 2-5-1, Minato-ku
Tokyo, 105-6218, Japan
tmuraya@ristex.jst.go.jp

Abstract

In these days, a main style with which access to the information on the network community is reading text in front of a personal computer and a mobiles. We designed, developed and evaluate new access style to the information on network community using story-based communication and presentation with conversation style. SPOC (Stream-oriented Public Opinion Channel) we developed provides a web-based multimedia environment aiming at helping non-skillful people to convert their stories into TV-like programs very easily. In this paper we propose a component of SPOC, S-RC (SPOC - RSS (RDF Site Summary) Compatibility), which enables sharing and cooperation of SPOC and the several network communities on CMS.

Keywords: inter face, agent, CMS, RSS, community

1 Introduction

By development of network technology, people can communicate each other over the network, exchange and offer the information using mailing list, Bulletin Board System (BBS), etc. These network applications (mailing list, BBC, etc) give the place to activity of network community, each community members exchange the information and share the knowledge in the community. Thus, this type of interaction has potential for creating and reconstructing knowledge in the community. As a method facilitating this process, we have proposed Social Intelligence Design for the network community (Nishida 2002). We designed, developed and evaluate new access and share style to the information and knowledge on network community using story-based communication. We developed a multimedia environment, SPOC (Stream Public Opinion Channel), which allows non-skillful users to (1) create and edit multimedia story content like a TV program and presentation, (2) publish the program on the network, and (3) play programs published by other community members. SPOC provides a media technology which automatically embodies a story using speech, graphics, movie, and animated agent, and supports story-based communication in a network community.

In this Paper, we propose cooperation with CMS (Content Management System) because we will take SPOC system more practice for especially people of outside of the laboratory. CMS is web application with site management functions, and a portal site and a business site can be built and managed easily, many community sites are managed on CMS. There is another advantage of CMS except assisting management practice and it takes information reorganization achieves the various purposes easy among each sites and each applications by using RSS(RDF Site Summary), as examples, the latest information of other sites is embedded and displayed on a site, tool which collects a round and the newest information for the site which publish favorite information automatically, the tool with which the newest information is displayed by flowing from the electrical scoreboard on PC ... etc, the information generated from the community site on CMS is sent to a user by various channels and various new applications. We think SPOC

contents to be used by people of outside of the laboratory as one of these channels and SPOC contents are sent to a user by another various channels by compatibility between SPOC and RSS. It enables user to make, enjoy and share information by using SPOC as one of channels from CMS, and we are trying that SPOC-RSS cooperation takes opportunity of using SPOC system increasing. We will describe at section 2 and 3, SPOC system and previous system, at section 4, various reorganization by CMS and RSS, at section 5, describe about compatibility between SPOC and RSS and at section 6, conclusion and future works.

2 Background

Basic idea and Concept of SPOC (Stream Public Opinion Channel), (Murayana et al. 2003) are inherited from POC (Public Opinion Channel) (Fukuhara et al. 2003) and EgoChat (Kubota et al. 2002). POC is a participatory broadcasting system that community member casts their information (question,and opinion,etc),the information contributed is expressed by primitive agent and picture, and it broadcast towards the community. By extending this system, Kubota (2002) developed EgoChat, which employs an agent-based approach in order to facilitate conversational information circulating process within a community. In EgoChat system, an embodied agent that acts as a virtualized ego talks on the user's behalf. The user can have a conversation with each virtualized user's personality on the system using speech interface.

3 SPOC

SPOC is a Web application, which anyone who has the internet access can use to create, edit, post, and play a program without a complicated software installation. All the materials are delivered from server side .The system overview is shown in Fig. 1. In SPOC, user' story is embodied as a SPOC program, which is like a TV program. A SPOC program consists of a sequence of Knowledge Cards as shown in Fig. 1. Each Card is like a scene in a TV program, and a user edits a Card one by one. Card is a building block composing a story. Knowledge card consists of text and a picture or a movie clip. Thus, users can make different stories by changing the order of the Cards. User's Stories is expressed by using embodied interface agent and multimedia. Although the expression by the interface agent requires detailed planning to control agent, in SPOC, an agent's planning of operation can be automatically created from a text by the CAST system (Nakano et al. 2004). By the SPOC system, a user only creates and rearranges a knowledge card and can publish contents like a TV program at a community very easily. Evaluation experiments (Nakano et al. 2004) showed that is easy to use in creating a program.

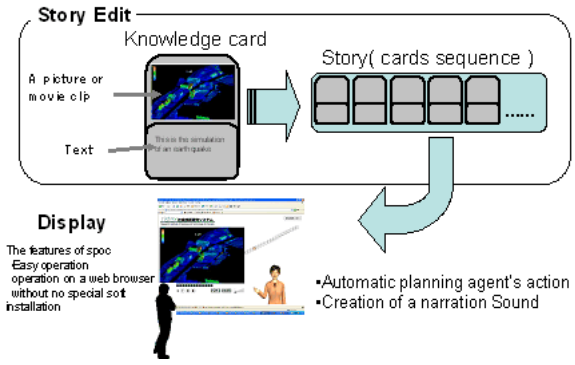


Figure 1: Overview of SPOC

4 CMS and RSS

The information created from the net community sites tended to be closed in the community site, in these days, the information is reorganized by various purposes and tools, and it is sent

to users by various channels and styles. CMS (Content Management systems) (see references Xoops and Zope) and RSS (RDF Site Summary) (RDF, 2000) have played the important role in community sites. CMS is web application with site management functions, and a portal site and a business site can be built and managed easily, many community sites are managed on CMS. In CMS, a data layer and an expression layer can be separated clearly in management of community sites, therefore, it is easy to transfer and reorganize the data of community sites. RSS is used for the standard of these data sharing among different web systems and tools. RSS is the format of the XML base which describes metadata, such as a title of a web site, and a summary. Individual web diaries, community sites and news sites, which release their latest information by RSS, have also been increasing. Fig. 2 shows that RSS takes information reorganization easy for each user's purposes. Some may collect information automatically and efficiently from some community sites, some may analyze huge information collected, some may check headline by like an electrical scoreboard, some may embedded information published by other community sites with relation at their community site. Separation of data by CMS and standardization of data by RSS provide various use styles of information from community sites. We try that SPOC contents are to be available as one of these styles.

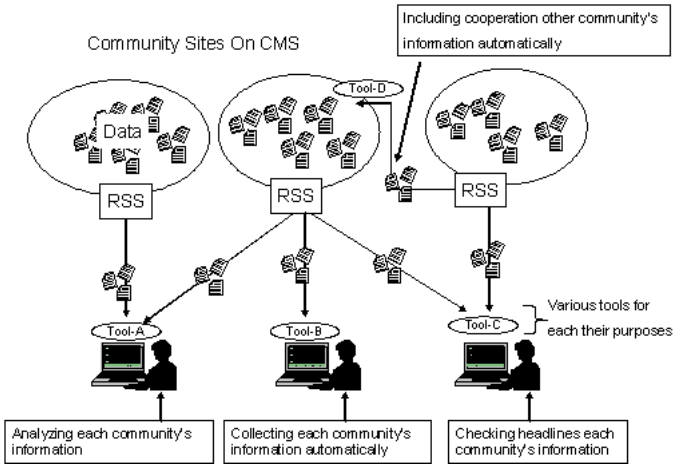


Figure 2: Overview of CMS and RSS usage

5 The Cooperation of SPOC, CMS, and RSS

5.1 S-RC (SPOC - RSS Compatibility) for the Cooperation

We describe cooperation of SPOC contents, CMS, and RSS. The overview of cooperation with CMS and SPOC is shown in Fig 3. The merit by the cooperation is that a user can use informational expression styles properly on a community site. Evaluation experiment (Nakano et al. 2004) showed SPOC conveys information comprehensibly rather than correctly. If a user wants to know and offer comprehensive information, SPOC contents will be helpful. It is SPOC - RSS Compatibility (S-RC) for a user to enable viewing, listening and creating of SPOC contents seamlessly on CMS. The cooperation of SPOC and CMS is performed as follows.

① CMS user publishes their knowledge, opinion and comment by text attached picture and movie from General web form CMS provides, and information published is listed by RSS format.

② SPOC imports the RSS from CMS information. Fig 4 showed equivalent between RSS and SPOC contents. Usually, the minimum unit of the user's knowledge, opinion and comments on RSS is expressed by "item" elements, and S-RC component relates "item" to "knowledge card" on SPOC. The "item" element has "description" element and "link" element, and

S-RC relates the "description" to text of knowledge card and the "link" to a picture and a movie clip of knowledge card. CMS information transferred from RSS to SPOC data format and SPOC database.

③ User can edit and modify SPOC contents from CMS and creates SPOC contents.

④ SPOC contents is published by RSS, CMS imports RSS on SPOC and CMS includes and displays SPOC contents to CMS users.

⑤ If CMS user wants to know and offer comprehensive information, the user can view and hear SPOC contents on CMS

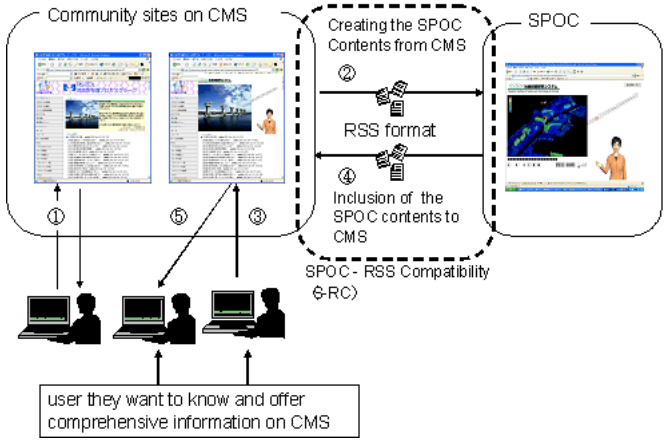


Figure 3: Overview of CMS and RSS usage

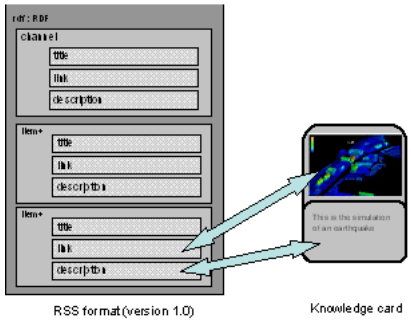


Figure 4: Overview of CMS and RSS usage

5.2 The Implemetne of the S-RC

Fig 5 shows the snapshot of cooperation of SPOC and RSS by using implemented SR-C. XOOPS (see reference Xoops) was used from some CMS soft wares. The S-RC consists of ① CMS's component for publishing CMS contents by RSS format, ② CMS's component for importing RSS format. ③ SPOC's component for importing RSS format, ④ SPOC's component for publishing SPOC contents by RSS format. Since many CMS and are fundamentally equipped with the functions of exporting and importing by RSS format, basically, implements of component ① and ④ are unnecessary. We implemented ② and ③ mainly. These coded functions are (1) shift to SPOC data base from XML of RSS form,(2)shift to RSS form from SPOC database fundamentally, and coded by JAVA language. User on CMS can transfer CMS contents to SPOC contents, edit SPOC contents and publish SPOC contents on CMS.

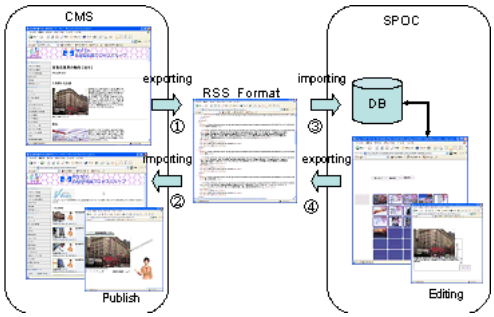


Figure 5: Overview of CMS and RSS usage

6 Conclusion and Future works

By spread of common XML formats to express information on community web sites, the information on a community web site is beginning to circulate seamlessly among various community web sites. In this paper we propose and implemented a component of SPOC - RSS Compatibility, which enables sharing of contents between SPOC and the several network communities on the web application by RSS. The S-RC provides function that user can transfer CMS contents to SPOC contents and publish SPOC contents on CMS, user is available to view CMS contents and SPOC content that conveys information comprehensibly rather than correctly. SPOC provides a media technology which automatically embodies a story using speech, graphics, movie, and animated agent, and supports story-based communication, and we conducted the evaluation experiment to the contents makers (Nakano et al. 2004) of SPOC. Now we are planning the evaluation experiment to contents watchers for evaluating the difference from other presentation software and the effect by the interface agent. Moreover, the collaboration experiment with community web site focusing on study of business administration community site is advanced using implemented SPOC - RSS Compatibility function.

References

Fukuhara, T., et al.(2003). A network-based interactive broadcasting system for supporting a knowledge-creating community, In *Internet-Based Intelligent Information Processing Systems*, R.J.Howlett et al., (Eds), 2003, World Scientific Publishing. p. 227-268.

Kubota, H., S. Kurohashi, and T. Nishida (2002). Virtualized-egos using Knowledge Cards. in Seventh Pacific Rim International Conference on Artificial Intelligence(PRICAI-02) WS-5 International Workshop on Intelligent Media Technology for Communicative Reality (IMTCR2002). 2002. Tokyo, JAPAN.

Murayama T., Y. Nakano., T.Nishida (2003). Participatory Broadcasting System Using Interface Agent and Multimedia, Social Intelligence Design International Conference, 2003.

Nakano Y., T.Murayama., T.Nishida (2004). Multimodal Story-based Communication:Integrating a Movie and a Conversational Agent, IEICE Transactions, 2004 (to appear).

Nishida, T. (2002). Social Intelligence Design for Web Intelligence, Special Issue on Web Intelligence. IEEE Computer, 2002. Vol. 35(No. 11): p. 37-41.

RDF Site Summary (RSS) 1.0. `http://web.resource.org/rss/1.0/`, 2000.

Xoops. `http://www.xoops.org/modules/news/`

Zope. `http://www.zope.org/`

Experimental Investigation into Influence of Negative Attitudes toward Robots on Human–Robot Interaction *

Tatsuya Nomura [1,2] †
[1]Department of Media Informatics, Ryukoku University
1–5, Yokotani, Setaohe–cho, Otsu, Shiga 520–2194, Japan
nomura@rins.ryukoku.ac.jp

Takayuki Kanda [2]
[2]Department of Communication Robots,
ATR Intelligent Robotics and Communication Laboratories
2–2, Hikaridai, Seika–cho, Soraku–gun, Kyoto 619–0288, Japan
kanda@atr.jp

Tomohiro Suzuki [3]
[3]Graduate School of Sociology, Toyo University
5–28–20, Hakusan, Bunkyo–ku, Tokyo 112–8606, Japan
suzukirt@h9.dion.ne.jp

Abstract

Negative attitudes toward robots are considered as one of psychological factors preventing humans from interacting with robots in daily life. To verify their influence on humans' behaviors toward robots, we designed and executed experiments where subjects interacted with Robovie, which is being developed as a platform for research on the possibility of communication robots. This paper reports and discusses the results of these experiments on correlation between subjects' negative attitudes and their behaviors toward robots. Moreover, it discusses influences of genders and experiences of real robots on their negative attitudes and behaviors toward robots.

Keywords: Human–Robot Interaction, Negative Attitudes toward Robots, Psychological Experiments, Gender Difference

1 Introduction

A great deal of study has been performed recently on robots that feature functions for communicating with humans in daily life, i.e., communication robots. This research has many applications such as entertainment, education, psychiatry, and so on (Dautenhahn et al., 2002; Druin and Hendler, 2000). If communication robots are really applied to these fields, however, it should be carefully investigated how humans are mentally affected by them.

Computer anxiety, an anxious emotion that prevents users from using and learning about computers, has been studied in educational psychology as an important factor in education for

*This research was supported by the National Institute of Information and Communications Technology.
†We thank Prof. Masahiro Sakamoto of Bunkyo Gakuin University for providing us with the Japanese version of the Personal Report of Communication Apprehension (PRCA–24) as his form of cooperation with our research on constructing the Negative Attitude toward Robots Scale. Moreover, we thank Dr. Kensuke Kato in Osaka University for his cooperation in checking the content validity of Negative Attitude toward Robots Scale.

Table 1: All the questionnaire items in the NARS

No.	Questionnaire Items
1	I feel anxiety if robots really have their own emotions.
2	I surmise that something negative for humans happen when robots become more similar to humans.
3	I will be able to be relaxed if I interact with robots.*
4	I feel anxiety when I imagine that I may be employed and assigned to a workplace where robots should be used.
5	I will be familiar with robots if they have their own emotions.*
6	I am mentally healed when I see robots behaving affectively. *
7	I am left helpless even by hearing something on robots.
8	I am likely to bring shame on myself when I use robots in public.
9	The words "artificial intelligence" or "decision by robots" make me feel unpleasant.
10	Even standing in front of robots will strain me.
11	I surmise that extreme dependence on robots may cause something negative for humans in future.
12	I will feel nervous if I interact with robots.
13	I am afraid that robots may negatively influence children's mind.
14	I surmise that future societies may be dominated by robots.

(*inverse item)

computer literacy (Raub, 1981; Hirata, 1990). Thus, influence of communication robots on children in pedagogical applications and clients in psychiatric applications should be also considered. This influence also should be clarified from perspectives of designs for communication robots in other daily–life applications.

Our research on this subject focuses on attitudes toward communication robots as an psychological construct. We consider negative attitudes toward robots as a psychological factor preventing individuals from interaction with robots having functions of communication in daily life, and have been developing a psychological scale measuring it, the Negative Attitude toward Robots Scale (NARS). [1] By using this psychological scale, we designed and executed experiments where subjects interacted with a humanoid type communication robot "Robovie," which is being developed as a platform for research on the possibility of communication robots (Ishiguro et al., 2003), to investigate the influence of their negative attitudes toward robots into their behaviors toward them.

This paper presents procedures and results of the human–robot interaction experiments, and discusses relations between negative attitudes and human behaviors toward communication robots. Moreover, we consider influences of gender difference and experiences of real robots on them.

2 Negative Attitude toward Robots Scale

The Negative Attitude toward Robots Scale (NARS) has been developed for measuring humans' attitudes toward communication robots in daily life. We have already confirmed its internal consistency and factorial validity (Nomura et al., 2004). In this paper, we mention only the overview of this confirmation process.

First, 32 candidates of questionnaire items were extracted from the freely described sentences in the pilot survey and the conventional psychological scales on computer anxiety and communication apprehension (Hirata, 1990; Pribyl et al., 1998), and their content validity was confirmed by two psychologists including one of the authors. Second, the pretest was executed based on these 32

[1] We tried to develop a psychological scale for measuring anxiety toward robots (Nomura and Kanda, 2003). After some analysis, it was clarified that our scale does not measure anxiety itself, but negative attitudes toward robots (Nomura et al., 2004).

Table 2: The subordinate scales and item numbers included in them in the NARS

Index	Subordinate Scales	Item No.
S1	Negative Attitude toward Situations of Interaction with Robots	4,7,8,9,10,12
S2	Negative Attitude toward Social Influence of Robots	1,2,11,13,14
S3	Negative Attitude toward Emotions in Interaction with Robots	3, 5, 6

items, and 263 data samples were assembled. Factor analysis and item analysis consisting of good–poor analysis, correlation coefficients, and α coefficients were executed for the pretest data, and as a result, 14 items included in 3 subordinate scales corresponding to 3 factors were extracted. Then, the test was executed based on these 14 items, and 240 data samples were assembled. It was confirmed by factor analysis that the test data had the factor structure consisting of 3 factors same as that in the pretest data. The indices of goodness–of–fit in this factor analysis were as follows: $GFI = 0.900$, $AGFI = 0.856$, $RMSEA = 0.080$. Moreover, α–coefficients of these 3 subordinate scales in the test data were 0.775, 0.782, and 0.648, respectively. At this time, its conceptual validity and test–retest reliability were also investigated (Nomura et al., 2004).

Table 1 shows the sentences of the questionnaire items obtained through the above confirmation process. Note that this scale has been developed in Japanese and the formal English version of it has not been completed. These sentences were roughly translated by one of the authors, not produced along formal procedures including back–translation. Moreover, table 2 shows these subordinate scales and item numbers included in them.

The number of grades in the answer at each item is five (1: I strongly disagree, 2: I disagree, 3: Undecided, 4: I agree, 5: I strongly agree), and the score of an individual at each subordinate scale is calculated by summing the scores of all the items included in the scale, with inverses of scores in some items. Thus, the minimum score and maximum score are 6 and 30 in the subordinate sale **S1**, 5 and 25 in the subordinate scale **S2**, and 3 and 15 in the subordinate scale **S3**, respectively.

3 Experimental Procedure

This section explains the robot, the Negative Attitude toward Robots Scale used as a controlled variable, and concrete procedures in our experiments of human–robot interaction.

3.1 Robovie

As shown in Figure. 1, Robovie is a robot that has a human–like appearance and is designed for communication with humans (Ishiguro et al., 2003). It stands 120 cm tall, its diameter is 40 cm, and it weights about 40 kg. The robot has two arms (4×2 DOF), a head (3 DOF), two eyes (2×2 DOF for gaze control), and a mobile platform (two driving wheels and one free wheel).

The robot has various sensors, including skin sensors covering the whole body, 10 tactile sensors located around the mobile platform, an omni–directional vision sensor, two microphones to listen to human voices, and 24 ultra–sonic sensors for detecting obstacles. It carries a Pentium III PC on board for processing sensory data and generating gestures. The operating system is Linux.

3.2 Procedures of Experiments on Human–Robot Interaction

Our experiments on human–robot interaction were executed in the room shown in Figure 1. Robovie programmed in advance was prepared for interaction with subjects in the room, and each subject communicated with it for a few minutes alone. The procedures used in one session of the experiments are shown as follows:

1. Before entering the experiment room shown in Figure 1, the subjects responded for the following questionnaire items:

   - sex, age,

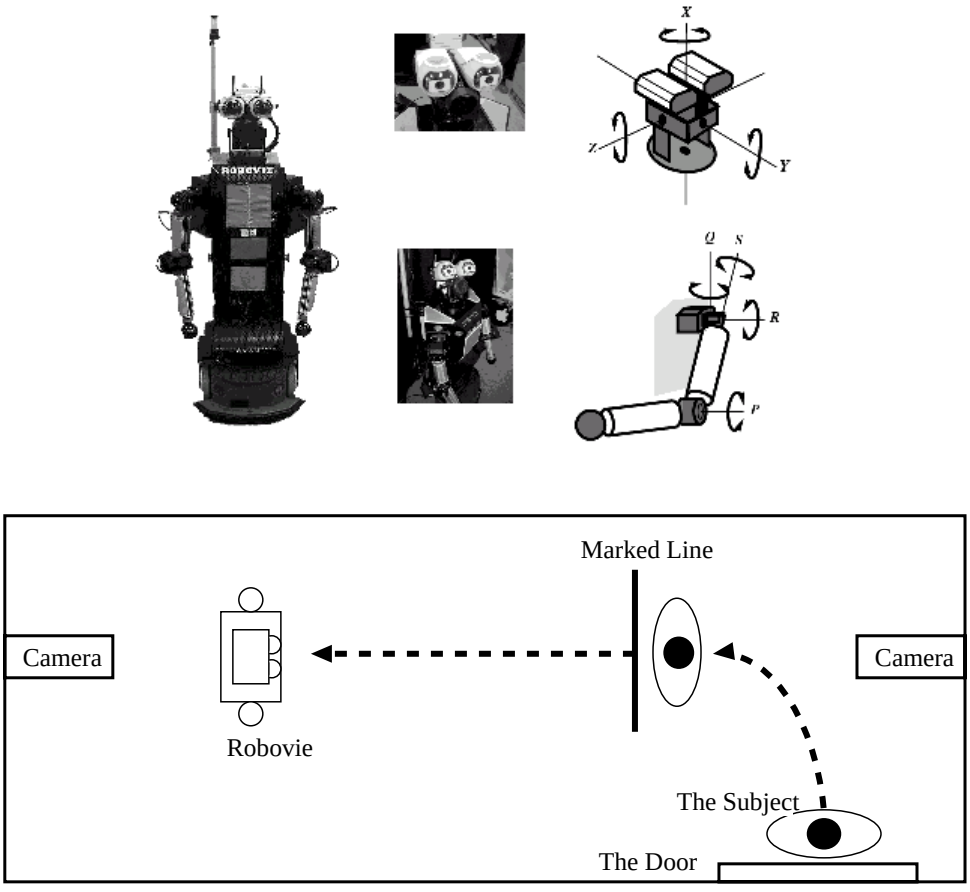


Figure 1: Robovie and the overview of the room where the experiments were executed (a view from above)

- whether he/she had seen really acting robots,
- the NARS.

2. Just before entering the room, they were instructed to talk toward Robovie just after entering the room.
3. The subject entered the room alone. Then, he/she moved to the marked line on the floor.
4. After he/she talked to Robovie, or a constant time (30 seconds) passed, Robovie uttered a sentence to stimulate his/her self–expression ("Have you recently experienced something negative?")
5. After he/she replied to the utterance of Robovie, or a constant time (30 seconds) passed, Robovie uttered a sentence to stimulate his/her physical contact to it ("Touch me").
6. After he/she touched the body of Robovie, or a constant time (30 seconds) passed, the session finished.

Behaviors of the subjects, including their utterances, were recorded using two digital video cameras as shown in Figure 1. Then, the following items related to their behaviors were extracted from the video data:

- The distance from the subjects to Robvie when they first stood in front of the robot after entering the room (**D**)

Table 3: The average values of the behavioral indices between the higher– and lower–score subgroups on each subordinate scale and t–values of the sores between the subgroups (H: the higher–score subgroup, L: the lower–score subgroup, n: the number of subjects in the subgroup. The values in brackets represent the standard deviation).

		D (mm)			**U1** (sec)	
		Average (SD)	t–Value		Average (SD)	t–Value
S1	H (n=25)	1264.9 (525.1)	1.01	H (n=21)	6.5 (1.6)	2.65*
	L (n=24)	1127.5 (392.6)		L (n=28)	5.2 (1.7)	
S2	H (n=25)	1161.1 (395.3)	-0.55	H (n=24)	5.6 (1.5)	-0.93
	L (n=24)	1235.6 (534.2)		L (n=25)	6.1 (2.0)	
S3	H (n=21)	1255.8 (507.3)	0.74	H (n=20)	6.2 (1.8)	1.02
	L (n=28)	1154.0 (434.9)		L (n=29)	5.6 (1.8)	
		U2 (sec)			**T** (sec)	
		Average (SD)	t–Value		Average (SD)	t–Value
S1	H (n=25)	2.4 (2.5)	0.28	H (n=22)	4.2 (3.4)	0.13
	L (n=22)	2.3 (1.4)		L (n=26)	4.1 (2.0)	
S2	H (n=24)	2.3 (1.3)	-0.17	H (n=25)	3.7 (1.8)	-1.17
	L (n=23)	2.4 (2.6)		L (n=23)	4.6 (3.3)	
S3	H (n=19)	2.4 (1.4)	0.22	H (n=20)	4.0 (2.6)	-0.32
	L (n=28)	2.3 (2.4)		L (n=28)	4.2 (2.8)	

($^*p < .05$)

- The time elapsed until the subjects talked to Robovie after entering the room (**U1**)
- The time elapsed until the subjects replied to Robovie after it uttered to stimulate their self-expression (**U2**)
- The time elapsed until the subjects touched the robot's body after it uttered to stimulate the subjects' physical contact with it (**T**)

Moreover, the contents of the subjects' utterances in the above step 5, that is, their replies to stimulation from the robot for their self–expression, were classified into three categories: utterances about something related to the subjects themselves, utterances about something not related to themselves, and non–utterance. This classification was executed by two persons, and if there was a difference between classification results of the two persons they discussed and integrated their classification results.

4 Experimental Results

This section shows results of the experiments shown in section 3. Fifty–three university students were asked to participate in the experiments as subjects (male: 22, female: 31), and the average age of these subjects was 19.9 (male: 20.6, female: 19.5). This paper focuses on the influence of the subjects' negative attitudes toward robots on their behaviors toward the robot and differences shown between genders. Moreover, influence of the subjects' experiences of real robots is considered.

4.1 Influence of Negative Attitudes on Behaviors toward the Robot

In order to clarify influence of the subjects' negative attitudes toward robots on their behaviors, we divided the subjects into two subgroups based on the median value of the scores of each subordinate scale of the NARS, then executed a t–test to verify a statistically significant difference on the behavior indices shown in section 3.2 between the subgroups at each subordinate scale. Table 3

Table 4: The average values of the NARS scores in the subgroups based on the contents of the subjects' utterances and f–values of the one–way ANOVA (G_1: the subgroup of the subjects who uttered about something related to themselves, G_2: the subgroup of the subjects who uttered about something not related to themselves, G_3: the subgroup of the subjects who did not utter anything, n: the number of subjects in the subgroup. The values in brackets represent the standard deviation).

	Average (SD)			f–Value	post–hoc test
	G_1 (n=9)	G_2 (n=39)	G_3 (n=3)		
S1	13.7 (4.0)	11.6 (3.1)	18.3 (1.2)	6.523^{**}	$G_2 < G_3{}^{**}$
S2	16.6 (3.4)	15.3 (3.2)	21.0 (1.6)	4.618^{*}	$G_2 < G_3{}^{*}$
S3	9.9 (2.3)	9.0 (2.3)	11.7 (2.1)	2.166	

($^{*}p < .05$, $^{**}p < .01$)

shows the average values of the behavior indices between the higher– and lower–score subgroups on each subordinate scale, and t–values of the sores between the subgroups. Note that there are differences in the number of subjects in the subgroups, dependent on the median values of the subordinate scales and indices since some behavior indices were not displayed by some subjects (the indeices **U1**, **U2**, and **T** were treated as lost data in case that the subjects did not respond within 30 seconds).

There was no statistically significant difference in the behavior indices **D**, **U2**, and **T** between the higher– and lower–score subgroups divided by any of the subordinate scales. However, there was a statistically significant difference of 5% on the time elapsed until the subjects talked to Robovie after entering the room (**U1**) between the higher– and lower–score subgroups based on the subordinate scale of negative attitude toward situations of interaction with robots (**S1**). This result suggests a possibility that persons with higher negative attitude toward situations of interaction with robots tend to avoid to talk to robots.

Moreover, we divided the subjects into three subgroups based on the contents of their utterances mentioned in section 3.2, then executed a one–way ANOVA with Tukey post–hoc tests to verify a statistically significant difference on the NARS scores between the subgroups. Table 4 shows the average values of the NARS scores in the subgroups and f–values of the ANOVA.

On the scores of negative attitudes toward situations of interaction with robots (**S1**) and social influence of robots (**S2**), there were statistically significant differences of 1% and 5% respectively. Moreover, it was found by the post–hoc tests that the scores of **S1** and **S2** in the subgroup of the subjects who did not utter anything were higher than those in the subgroup of the subjects who uttered something not related to themselves with statistical significance of 1% and 5%, respectively. This result suggests that persons with higher negative attitudes toward situations of interaction with robots and social influence of robots tend to avoid their self–expression to robots.

4.2 Influence of Gender on Relations between Negative Attitudes and Behaviors toward the Robot

First, we executed a t–test to verify statistically significant differences in the scores of NARS and behavior indices between the male and female subjects. Table 5 and 6 show the average values of the NARS scores and behavior indices, and t–values of them between the males and females.

As shown in Table 5, there was a trend that the female subjects had lower negative attitudes toward robots than the male subjects. In particular, there was a statistically significant difference of 5% on the scores of negative attitude toward emotions in interaction with robots (**S3**). Moreover, as shown in Table 6, there was a statistically significant difference of 0.1% for the distance from the subjects to the robot when they first stood in front of the robot after entering the room (**D**).

Second, we investigated correlation coefficients between the NARS scores and behavior indices independently for male and female subjects. Table 7 shows these correlation coefficients.

Table 5: The average values of the NARS scores in the male and female subjects and t–values between them (n: the number of subjects. The values in brackets represent the standard deviation).

	S1		**S2**		**S3**	
	Average (SD)	t–Value	Average (SD)	t–Value	Average (SD)	t–Value
Males (n=22)	12.6 (3.9)	0.455	16.3 (3.2)	0.751	10.1 (2.0)	2.267*
Females (n=31)	12.1 (3.3)		15.5 (3.6)		8.6 (2.4)	

($^{*}p < .05$)

Table 6: The average values of the behavior indices in the male and female subjects and t–values between them (n: the number of subjects. The values in brackets represent the standard deviation).

		D			**U1**	
	n	Average (SD)	t–Value	n	Average (SD)	t–Value
Males	21	1479.3 (513.4)	3.860***	19	5.7 (1.9)	-0.410
Females	28	986.3 (291.1)		30	5.9 (1.7)	
		U2			**T**	
	n	Average (SD)	t–Value	n	Average (SD)	t–Value
Males	18	2.2 (1.3)	-0.445	20	4.9 (3.3)	1.530
Females	29	2.4 (2.4)		28	3.6 (2.0)	

($^{***}p < .001$)

On the time elapsed until the subjects talked to the robot after entering the room (**U1**) and their scores of negative attitude toward emotions in interaction with robots (**S3**), the female subjects showed a statistically significant positive correlation to a medium level, whereas the male subjects showed a low correlation. Although there was no statistical significance, on the time elapsed until the subjects replied to the robot after it uttered to stimulate their self-expression (**U2**) and their scores of negative attitude toward situations of interaction with robots (**S1**), the male subjects showed a negative correlation whereas the female subjects showed a positive correlation. Moreover, on the distance from the subjects to the robot when they first stood in front of it after entering the room (**D**) and their scores of negative attitude toward emotions in interaction with robots (**S3**), on which there was a statistically significant difference between the male and female subjects in Table 5, the male subjects showed a positive correlation, whereas the female subjects showed a negative correlation. On the time elapsed until the subjects talked to the robot after entering the room (**U1**) and their scores of negative attitude toward social influence of

Table 7: Correlation coefficients between the NARS scores and behavior indices in the male and female subjects.

		D	**U1**	**U2**	**T**
S1	Males	0.162	0.210	-0.351	0.070
	Females	-0.022	0.141	0.260	-0.014
S2	Males	0.232	0.387	-0.244	0.245
	Females	-0.057	0.015	-0.044	-0.025
S3	Males	0.267	-0.057	-0.112	-0.115
	Females	-0.139	0.325†	0.040	-0.292

($^{\dagger}p < .1$)

Table 8: The average values of the NARS scores in the subgroups of the subjects who had seen really acting robots and those who had not, and t–values between them (n: the number of subjects, EE: the subgroups of the subjects who had seen really acting robots, NEE: the subgroups of the subjects who had not seen really acting robots. The values in brackets represent the standard deviation).

	S1		**S2**		**S3**	
	Average (SD)	t–Value	Average (SD)	t–Value	Average (SD)	t–Value
EE (n=20)	11.6 (3.5)	-1.216	16.1 (3.3)	0.410	9.1 (2.5)	-0.343
NEE (n=33)	12.8 (3.5)		15.7 (3.5)		9.3 (2.2)	

Table 9: The average values of the behavior indices in the subgroups of the subjects who had seen really acting robots and those who had not, and t–values between them (n: the number of subjects, EE: the subgroups of the subjects who had seen really acting robots, NEE: the subgroups of the subjects who had not seen really acting robots. The values in brackets represent the standard deviation).

		D			**U1**	
	n	Average (SD)	t–Value	n	Average (SD)	t–Value
EE	18	1352.1 (450.3)	$1.174^{\dagger}$	18	5.9 (1.8)	0.192
NEE	31	1107.9 (457.7)		31	5.8 (1.8)	
		U2			**T**	
	n	Average (SD)	t–Value	n	Average (SD)	t–Value
EE	17	3.2 (2.6)	$1.902^{\dagger}$	19	4.6 (2.5)	1.022
NEE	30	1.9 (1.4)		29	3.8 (2.8)	

($^{\dagger}p < .1$)

robots (**S2**), the male subjects showed a medium positive correlation, whereas the female subjects showed a low correlation. These values do suggest a possibility of gender differences in both negative attitudes toward robots and relations between them and behaviors toward robots.

4.3 Influence of Experiences of Real Robots on Relations between Negative Attitudes and Behaviors toward the Robot

As mentioned in section 3.2, the subjects responded for a questionnaire item asking whether they had previously seen really acting robots. We divided the subjects into the subgroups of those who denied it and those who acknowledged it, then executed a t–test to verify statistically significant differences in the scores of NARS and behavior indices between these subgroups. Table 8 and 9 show the average values of the NARS scores and behavior indices, and t–values of them between these subgroups.

As shown in Table 8, there was no statistically significant difference on the NARS scores between the subgroups of the subjects who had seen really acting robots and those who had not. However, as shown in Table 9, there were statistically significant tendencies of 10% on the distance from the subjects to the robot when they first stood in front of the robot after entering the room (**D**) and the time elapsed until the subjects replied to the robot after it uttered to stimulate their self-expression (**U2**).

Second, we investigated correlation coefficients between the NARS scores and behavior indices independently for the subgroups of the subjects who had seen really acting robots and those who had not. Table 10 shows these correlation coefficients.

On the time elapsed until the subjects replied to the robot after it uttered to stimulate their self-expression (**U2**) and their scores of negative attitude toward situations of interaction with robots

Table 10: Correlation coefficients between the NARS scores and behavior indices in the subgroups of the subjects who had seen really acting robots and those who had not (EE: the subgroups of the subjects who had seen really acting robots, NEE: the subgroups of the subjects who had not seen really acting robots).

		D	**U1**	**U2**	**T**
S1	EE	0.060	0.119	$0.478^{\dagger}$	0.140
	NEE	0.198	0.213	0.048	0.061
S2	EE	-0.075	0.256	-0.059	0.355
	NEE	-0.069	0.131	-0.103	0.109
S3	EE	0.144	0.058	-0.112	0.035
	NEE	0.258	0.216	0.195	-0.192

($^{\dagger}p < .1$)

(**S1**), the subjects who had seen really acting robots showed a medium positive correlation with statistically significant tendency of 10% whereas the subjects who had not seen robots showed a low correlation. Although there was no statistical significance, there was a similar tendency on the time elapsed until the subjects touched the robot's body after it uttered to stimulate the subjects' physical contact with it (**T**) and their scores of negative attitude toward social influence of robots (**S2**). This result suggests a possibility that individuals' experiences of real robots influence on relations between negative attitudes and behaviors toward robots.

5 Conclusion and Discussion

In this paper, we showed the procedures and results of our experiments on human–robot interaction by using a humanoid robot "Robovie" and the Negative Attitude for Robots Scale (NARS). As a result, we suggested a possibility that negative attitudes for robots affected human behaviors toward communication robots. Moreover, we noticed a possibility that there were gender differences in negative attitudes toward robots, and that there were also gender differences in relations between negative attitudes and behaviors toward robots. Furthermore, we noticed a possibility that individuals' experiences of real robots influence on relations between negative attitudes and behaviors toward robots.

The results of our experiments in section 4.1 show that negative attitude toward situations of interaction with robots affects interaction with communication robots, and this negative attitude and negative attitude toward social influence of robots affect self–expression toward communication robots. Mental disaffiliation is a common characteristic in behaviors associated with communication apprehension (Sakamoto et al., 1998), and the results suggest that persons with highly negative attitudes toward robots mentally tend to avoid human–robot communication.

Moreover, the results of our experiments in section 4.2 show that men and women differ in their degrees of negative attitudes toward robots, and correlation between the negative attitudes and communication behaviors such as utterances toward robots. The results in section 4.3 show that persons having seen really acting robots and those not having differ in correlation between the negative attitudes and communication behaviors such as utterances toward robots. These facts suggest that designs of communication robots' appearance and behaviors should be considered from the perspective of genders and individuals' experiences, in particular, in pedagogical and psychiatric fields.

However, our research has some problems.

We showed just a possiblity of influence of negative attitudes toward robots into behaviors toward them and gender difference in it, and the results of our experiments did sufficiently not clarify relations between the behavior indices and scores of NARS. In other words, there is a possibility that negative attitudes toward robots may directly not affect behaviors toward them. As a cause of it, we consider the fact that communication robots have been less popular than

computers, of which concrete images have been constructed in general people. Images of robots are wide from arm robots in factories to pet–type robots. Thus, it is guessed that psychological attitudes toward robots are hard to be connected with behaviors toward them in situations of real interaction with them, in comparison with computers. In order to predicate individuals' communication avoidance behaviors toward robots, we need to develop another psychological scale to measure anxiety or fear in situations of interaction with robots. Moreover, we should clarify which psychological mechanism causes influences by executing more detailed experiments and more complex statistical methods such as structural equation models.

Moreover, we need to solve the problem of sampling bias of subjects in this type of experiments. In other words, there is a difference on negative attitudes toward robots between subjects participating in human–robot interaction experiments and more general persons. We assembled data consisting of 238 respondents (male: 146, female 92, the average age of the male: 21.8, that of the females: 22.4) in order to investigate validity of the NARS (Nomura et al., 2004). In analysis of the data, it was found by a two–ways ANOVA that genders and experiences on robots affect the scores of the subordinate scale **S1** (see Table 11 in Appendix). The statistical trend in this data did not appear in the subjects of our experiment. Moreover, we executed a two–ways ANOVA for male–female and these respondents–the subjects in section 4 to investigate difference on the NARS scores between them. As a result, it was found that the female subjects had lower negative attitudes toward emotions in interaction with robots than the respondents (see Table 12 in Appendix. Tukey post–hoc tests confirmed it with statistically significance of 1%). This type of bias influences analysis and interpretation of human–robot interaction, and it is necessary to carefully consider it, in particular, when applying the results to the design of communication robots.

Furthermore, there is a possibility that negative attitudes toward communication robots, images of robots, and relations between the negative attitudes and behaviors differ dependent on cultures. By developing the English version of NARS and combining it with other psychological scales such as STAI (Hidano et al., 2000; Spielberger et al., 1970) and the robot image scale (Suzuki et al., 2002), we should investigate international comparisons of negative attitudes toward robots and their relations with behaviors toward and images of robots.

References

Dautenhahn, K., Bond, A. H., namero, L. C., and Edmonds, B. (2002). *Socially Intelligent Agents: Creating Relationships with Computers and Robots*. Kluwer Academic Publishers.

Druin, A. and Hendler, J. (2000). *Robots for Kids: Exploring New Technologies for Learning*. Morgan Kaufmann.

Hidano, T., Fukushima, M., Iwasaki, M., Soga, S., and Spielberger, C. D. (2000). *STAI Manual*. Jitsumu Kyoiku Shuppan. (in Japanese).

Hirata, K. (1990). The concept of computer anxiety and measurement of it. *Bulletin of Aichi University of Education*, 39:203–212. (in Japanese).

Ishiguro, H., Ono, T., Imai, M., and Kanda, T. (2003). Development of an interactive humanoid robot "Robovie": An interdisciplinary approach. In Jarvis, R. A. and Zelinsky, A., editors, *Robotics Research*, pages 179–191. Springer.

Nomura, T. and Kanda, T. (2003). On proposing the concept of robot anxiety and considering measurement of it. In *Proc. 12th IEEE International Workshop on Robot and Human Interactive Communication*, pages 373–378.

Nomura, T., Suzuki, T., Kanda, T., and Kato, K. (2004). The relationship between the attitude toward robots and the response for robots: Through an attempt for construction of Negative Attitude toward Robots Scale (NARS). *The Japanese Journal of Psychology*. (submitted).

Pribyl, C. B., Keaten, J. A., Sakamoto, M., and Koshikawa, F. (1998). Assessing the cross–cultural content validity of the Personal Report of Communication Apprehension scale (PRCA–24). *Japanese Psychological Research*, 40:47–53.

Raub, A. C. (1981). *Correlates of comuter anxiety in college students*. PhD thesis, University of Pennsylvania.

Sakamoto, M., Pribyl, C., and Keaten, J. (1998). The history and current state of communication avoidance research. *The Japanese Journal of Psychology*, 48:491–507. (in Japanese).

Spielberger, C. D., Gorsuch, R. L., and Lushene, R. E. (1970). *Manual for the State–Trait Anxiety Inventory*. Counseling Psychologist Press.

Suzuki, K., Kashibuchi, M., Sakamoto, A., and Nagata, J. (2002). Constructing an image scale for robots and investigating contents of measured images: Part I. In *Proc. 66th Annual Meeting of the Japanese Psychological Association*, page 114. (in Japanese).

Appendix

The following tables show the results of analysis for 238 respondents for the NARS assembled for investigating its validity in Nomura et al. (2004). Note that these results were not described in Nomura et al. (2004).

Table 11: The result of the two–ways ANOVA for the NARS scores in 238 respondents (n: the number of respondents, EE: the subgroups of the respondents who had seen really acting robots, NEE: the subgroups of the respondents who had not seen really acting robots).

	Average (SD)				f–Values		
	Males		Females		Factor 1	Factor 2	Mutual
	EE (n=124)	NEE (n=22)	EE (n=53)	NEE (n=39)	(Male–Female)	(EE–NEE)	Interaction
S1	10.7 (3.9)	12.0 (4.0)	11.7 (3.4)	14.1 (4.4)	8.997^{**}	6.993^{**}	0.785
S2	15.1 (4.3)	16.5 (4.1)	16.6 (4.2)	16.6 (4.9)	1.111	1.386	0.945
S3	10.5 (2.4)	10.0 (2.4)	10.5 (2.5)	10.4 (2.3)	0.507	0.224	0.248

($^{**}p < .01$)

Table 12: The result of the two–ways ANOVA for the NARS scores in the 238 respondents and the subjects in the human–robot interaction experiments (RE: the respondents for validity confirmation, SJ: the subjects in the human–robot interaction experiments).

	Average (SD)				f–Values		
	Males		Females		Factor 1	Factor 2	Mutual
	RE (n=146)	SJ (n=22)	RE (n=92)	SJ (n=31)	(Male–Female)	(RE–SJ)	Interaction
S1	10.8 (3.9)	12.6 (3.9)	12.8 (4.1)	12.1 (3.3)	1.398	0.848	$3.768^{\dagger}$
S2	15.3 (4.3)	16.3 (3.2)	16.6 (4.5)	15.5 (3.6)	0.188	0.006	2.367
S3	10.4 (2.4)	10.1 (2.0)	10.5 (2.4)	8.6 (2.4)	$3.375^{\dagger}$	8.610^{**}	4.022^{*}

($^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$)

A Method of Caring Interaction with Clients Based on a Nursing Theory and Naive Psychological Approaches

A. Notsu, O. Katai, H. Kawakami and T. Shiose
Department of Systems Science, Graduate School of Informatics
Kyoto University, Yoshida-Honmachi, Kyoto 606-8501, Japan
Phone:+81-75-753-3592, Fax:+81-75-753-5042
{notsu@sys.|katai@|kawakami@|shiose@}i.kyoto-u.ac.jp

Abstract

A framework of "caring interaction" via network representation of a client's concepts based on Nursing Theory and Naïve Psychological approaches is introduced. This network modeling of clients is based on a well-known theory of nursing, the Roy Adaptation model, as well as Heider's theory of Cognitive Balance from his Naïve Psychology. Heider's theory provides a method to analyze the cognitive consistencies ("cognitive balance") of the network model. This analysis is carried out by detecting cognitive inconsistencies systematically, which are then related to various "modes of caring interaction" via the Roy Adaptation model of nursing.

Keywords: Caring Interaction, Nursing Theory, Naïve Psychology and Narrative Therapy

1 Introduction

In this paper, we propose a framework of caring interaction based on a Nursing Theory and an Interpersonal Psychological approach referred to as Naive Psychology. The essential point here is the method of organizing various messages from clients and related pieces of knowledge on the field. In this paper, we adopted for this framework a well-known theory of nursing, the Roy Adaptation model, and Heider's theory on Cognitive Balance from his Naive Psychology. These bases provide a sophisticated way of organizing messages and pieces of knowledge to derive effective caring interactions with clients. In Section 2, a brief survey of nursing theories in shown, and then Roy's theory of nursing is introduced. In Section 3, Heider's theory of Naive Psychology and the notion of Perceptual (Cognitive) Balance are introduced. In Sections 4, these methods are integrated into a novel network modeling of a client's status that is suitable for effective Caring Interaction. In Section 5, we examine the modeling in practice and consider its results.

2 Nursing Theories and the Roy Adaptation Model

2.1 General Aspects of Nursing Theories

Nursing theories involve inventing and discovering various focal concepts and methods of nursing by examining events and their relevance in nursing interventions for the purposes of describing, explaining and prescribing caring processes. Such theories provide us with basic frameworks and aspects for organizing various pieces of knowledge, information and events in nursing processes (George, 1995). The monumental book "Notes on Nursing: What It Is and What It Is Not" by Florence Nightingale had not been discussed for many years. Since 1950, various theories motivated by the Notes (Nightingale, 1895) have been developed.

2.2 The Roy Adaptation Model

The "Roy Adaptation Model"(Roy, 1999; Andrews, 1986) of nursing regards a human being as an adaptation system. The clients' adaptive behavior is regarded as the product of four kinds of

"Adaptive Modes"(Fig. 1), i.e., Physiological Mode, Self-concept Mode, Role Function Mode and Interdependence Mode.

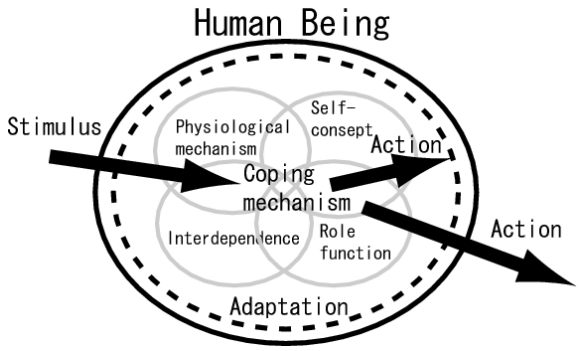


Figure 1: Human being as an adaptation system

Thus, environments are regarded as consisting of stimuli that are inputs to yield clients' responses in the Roy Adaptation Model. Nurses modify stimuli to promote the clients adaptation in these adaptation modes.

According to her definition, whether a client action is adaptive depends on whether it is effective for attaining the goal of the client such, as survival, development and maturation. These adaptation modes are regarded as adaptive response systems of human beings by processing various stimuli under various environmental changes.

3 Naive Psychology and Perceptual Balance

3.1 Naive Psychology and the Notion of Perceptual Balance

In the application of the nursing model described in Section 2 with the aid of a "Communication Message Board", we need a method to analyze the content of messages and the interrelationships among them. For this purpose, we refer to the Naive Psychology or Common-Sense Psychology proposed by F. Heider (Heider, 1958). He is famous for his Theory of Perceptual Balance or Cognitive Balance, usually referred to as P-O-X Theory (Harary, 1961)-(Newcomb, 1953).

He analyzed words that express the situation of a person ("Word Analysis") as well as the situation itself ("Situation Analysis") based on Naive Psychology. He arrived at the notion of "Relation Types" that underlie various situations as follows:

1. Relations of "experiencing or being affected"
2. Relations of "causing"
3. Relations of "can (being able to)"
4. Relations of "trying"
5. Relations of "wanting"
6. Relations of "belonging"
7. Relations of "ought and may"
8. "Sentiment" Relations

Relations (1)-(7), although not Sentiment Relations (8), are called "Unit Formation Relations". All of these relations are categorized as being either positive or negative, i.e., into (1)+, (1)-, (2)+, (2)-,..., (8)+ and (8)-.

He focused on the "consistency" of these relations in a very localized setting of situations, that is, "the perceptual or cognitive balance" of a person (noted as "p") with another person ("o") with regard to an entity ("x"). For example, let us consider the case where

- P likes o (we denote this as pLo: positive Sentiment Relation),
- O makes a record cabinet (x) (oUx: positive Unit Formation Relation),
- P thinks that the record cabinet is poorly made (pDLx: negative Sentiment Relation).

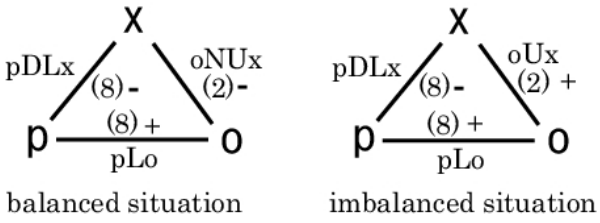


Figure 2: Heider's perceptually balanced and unbalanced situations

This situation is shown in the right side of Fig. 2 and is regarded as being "imbalanced" or "unbalanced". The balance of this triangular system is defined as the sign of the product of these three arcs being (+). In this case, we have (+)x(+)x(-) = (-), so the situation is imbalanced. The balanced situation can be accepted by p without "stress", while the imbalanced or unbalanced situation makes p feel stressful and uncomfortable. The theory of Perceptual Balance by Heider says that an imbalanced state (situation) is altered to restore the balance of the system (situation). Figure 3 depicts several possible relations between the three entities p, o and x, used to restore balance.

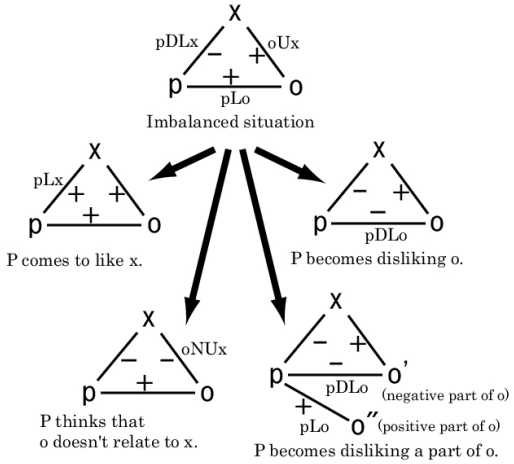


Figure 3: Restoration of balance from an imbalanced situation

4 Network Modeling of Clients

Based on the Roy Adaptation Model on nursing, we can expect enhancement of the Adaptation Mode, particularly in Self-concept Mode, Role function Mode and Interdependence Mode. The

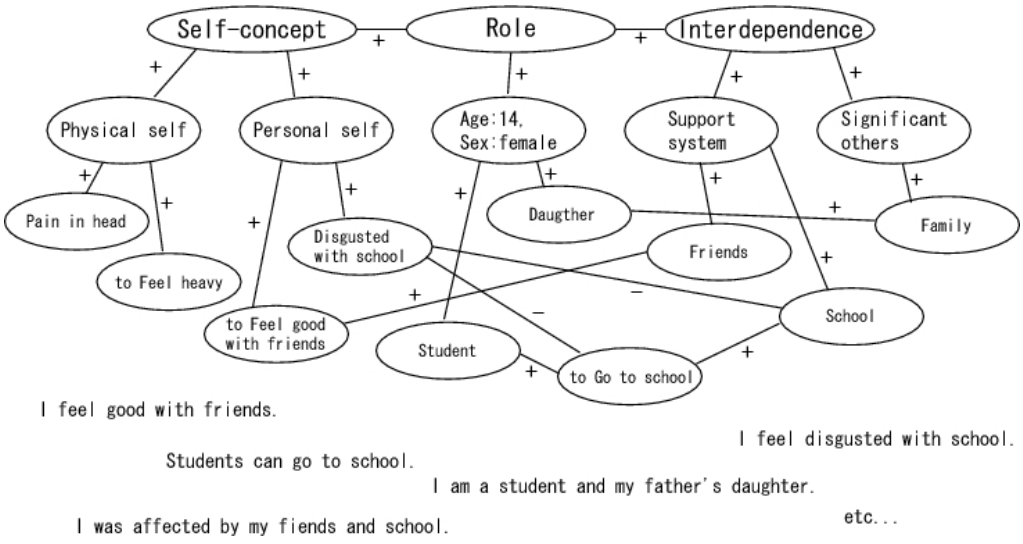


Figure 4: Network model of the concepts on a client

interesting point here is that the direction of enhancement or intervention can be considerably dependent on the relation type. Roughly speaking, we can readily see the following correspondence between the relation types (in Naive Psychology) and the basic notions in the Roy Adaptation Model (Notsu and Kawakami, 2001):

1. Relation of experiencing or being affected ⇔ Attention-seeking mode (Interdependence mode)Affection-seeking mode (Interdependent mode)
2. Relation of causing ⇔ Support-seeking mode (Interdependent mode)
3. Relation of can ⇔ Role function mode
4. Relation of trying ⇔ Role function mode
5. Relation of wanting ⇔ Self ideal mode (Self-concept)
6. Relation of belonging ⇔ Role function mode
7. Relation of ought and may ⇔ Moral-ethical self mode (Self-concept)
8. Sentiment Relation (sentiment message)⇔Self-consistency (Self-concept)

By the using these types, we can organize a network model of clients based of the framework of the Roy adaptation model.

On this network model, we incorporate "signs" ("+" or "-") to the links (arcs) by which we can regard models as "signed graphs". The basic links in the conceptualization of adaptation modes, such as self-concept mode, role mode and interdependence mode, are given such as the ones shown in Fig. 3. We can see that these links should be associated with the positive sign (+). This is because these links basically stand for "is-a" links (class-subclass links), that is, the upper nodes are more general concepts than the lower nodes or, equivalently, the lower nodes are more concrete than the upper ones (Katai and Iwai, 1978).

If the nodes are positively (negatively) related, the links are given with the positive (negative) sign "+" ("-"). This decision may be made by referring to the related social common sense or by the judgment of the clients to be modeled.

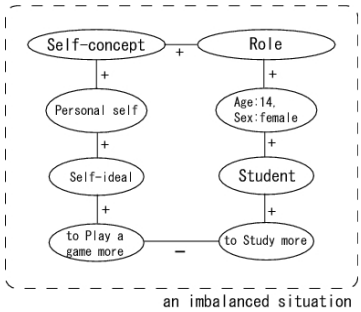


Figure 5: An imbalanced state (cycle)

4.1 Caring communication as balancing processes

These signed graphs (network) usually have complex structures and may involve imbalanced cycles. The existence of imbalanced cycles corresponds to the occurrence of a conflict that causes an adaptive behavior, which reduces the person's stress. For example, suppose that a person has said, "I want to play the game more. But I have to study more." In this case, we have the network structure shown in Fig. 5. In the figure, the area encircled by broken lines shows the "focused" area. That is, the person focuses himself (herself) on the area in his (her) mind. This area involves an imbalanced cycle, so the area is imbalanced. We need caring communications as balancing processes.

5 Example of Client Network Model

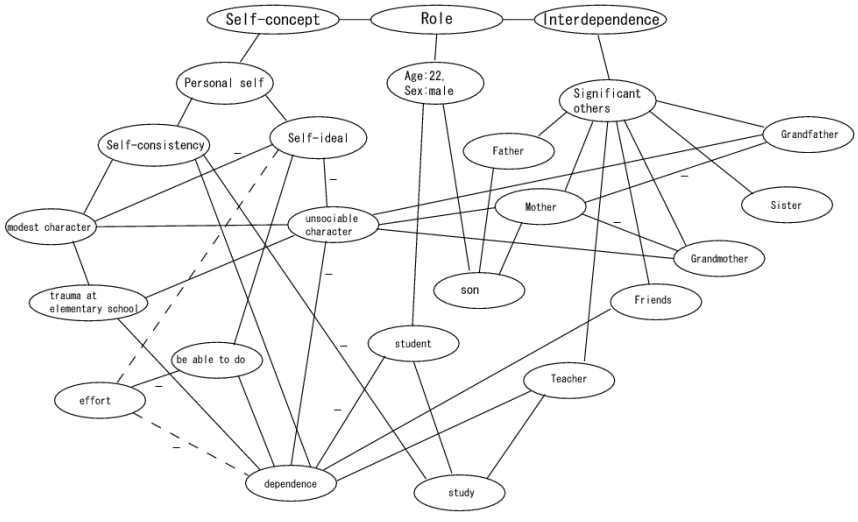


Figure 6: First example (Case 1) of client's concept network model

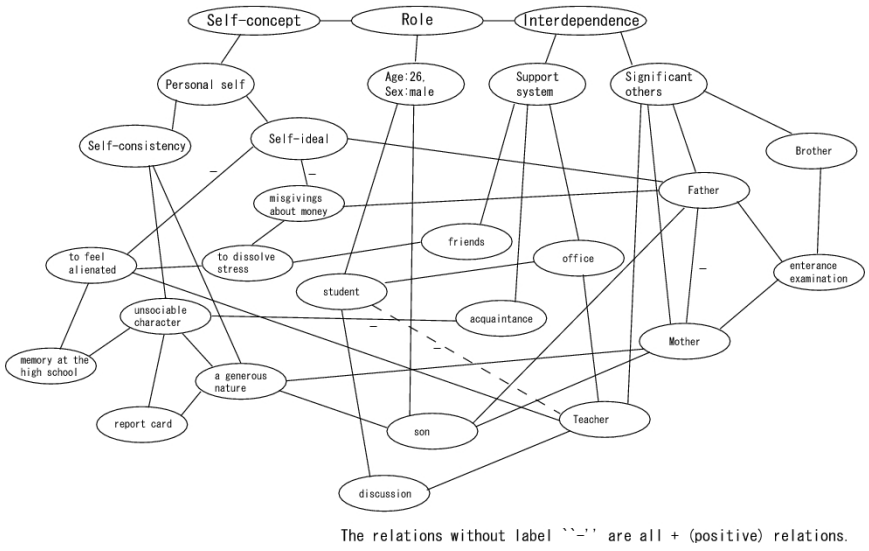


Figure 7: Second example (Case 2) of client's concept network model

We performed actual caring show clearly the validity of this conceptual network model and to see what obstacles pose actual problems. We clarified the meaning and the problems of this communication by referring to the narrative therapy approach in this chapter.

Figure 6 and 7 show Network Models of students that were made by talking with them. In this taking, there are some balancing processes used as caring communications. The following subsections show some of these balancing prccesses.

5.1 Change in network structure

Let's consider the follow pair of exchanges.

Horizontal shifting (Fig. 8)

A: "I thought that I should work harder. But I couldn't..."

B: "You have already worked hard. Don't worry about it."

A: "Now, I think I'll be OK."

B nods.

A: "So, now I feel better."

We know that he balanced his network structure by person A's "horizontal shifting" from the notion of "effort" to the notion of "I'm able to do it".

Incorporation of a new concept (fig. 9)

A: "I want to do another research project that is difficult in this lab."

B: "There may be another person who has the same problem."

A: "Hmm..."

B: "Anyway, you should discuss what to do with our teacher."

A: "You're right."

Person A balanced his network structure by incorporating a new concept of "discussion" .

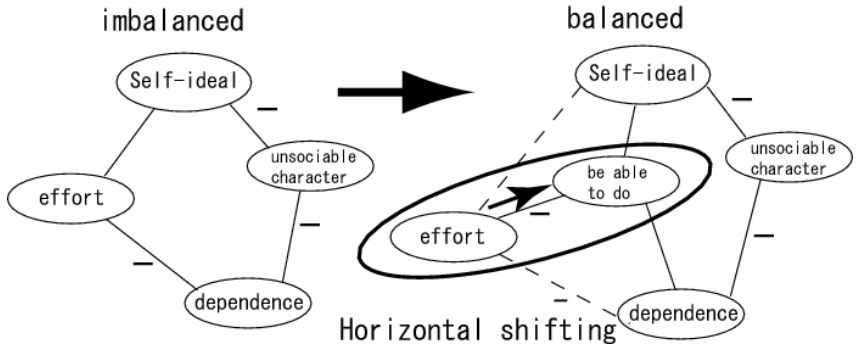


Figure 8: Horizontal shifting in Case 1

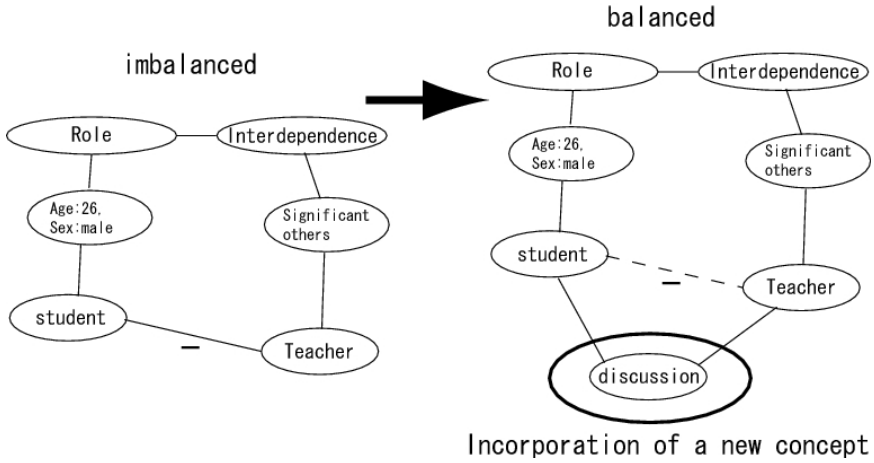


Figure 9: Incorporation of a new concept in Case 2

5.2 Dominant Story and Alternative Story

As the method of caring communication, we refer to the "Narrative Therapy" approach (McNamee and Gergen, 1992). In this feild, a therapist helps a client to change his/her "Dominant Story" into an "Alternative Story", where a"dominant story" is a socially consistent but stressful story to a person (client) while an "alternative story" is not, on the other hand, subject to commonly accepted social norms but less stressful to the person, and thus preferable to him/her. To change one balanced network into another balanced network means to change a dominant story into an alternative one, where the Self-Image (the network model) is substantiated as a "story along a series of experiences".

5.3 Consideration of Self-Narrative and Balance of network structure

In Case 1, the student might draw the formal network model in the following way.

He could be damaged by his friend's casual words during his an elementary school days, which are usually forgotten, but his sturucture was easily balanced by these words. This is to say, his structure got the notion "I am unsociable". Then, it is assumed that his network strcuture might be strengthened conversely by introducing other concepts. Also in his words, notions were unnaturally connected with those concepts. For example, this includes his saying "I am hopeless".

6 Conclusions

We have put a framework of caring interaction into practice. Our results show "caring" as a balancing process of a client's network structure and the effect on understanding another person. However, we should be careful in assessing and predicting the various possible effects of introducing this kind of interaction mechanism into practice. This is because there are many unconscious activities that occur in our mind.

References

Andrews H.A. and Sister Callista Roy (1986). *Essentials of the Roy Adaptation Model*, Appleton-Century-Crofts.

Cartwright D. and F. Harary (1956). Structural Balance, a Generalization of Heider's Theory, In *Psychological Review*, 63, pages 167–293.

George L.B., (ed.) (1995). *NURSING THEORIES: The Basis for Professional Nursing Practice*, Appleton & Lange.

Harary F. (1961). A Structural Analysis of Situation in the Middle East in 1961, In *J. Conflict Resolution, Vol. 5, No. 2*, pages 167–178.

Heider F. (1958). *The Psychology of Interpersonal Relations*, John Wiley.

Katai O. and S. Iwai (1978). Studies on the Balancing, the Minimal Balancing, and the Minimum Balancing Processes for Social Groups with Planar and Nonplanar Graph Structures, In *J. Math. Psychol., Vol. 18, No. 2*, pages 141–176.

Katai O. and S. Iwai (1978). On the Characterization of Balancing Processes of Social Systems and the Derivation of the Minimal Balancing Processes, In *IEEE Trans., Vol. SMC-8, No.5*, pages 337–348.

McNamee S. and K. J. Gergen (1992). *Therapy as Social Construction (Inquiries As Social Construction)*, Sage Publication.

Newcomb T.M. (1953). An Approach to the Study of Communicative Acts, In *Psychological Review*, 60, pages 393–404.

Nightingale, Florence N, (1895). *Note on nursing: What it is and what it is not (Com. Ed.)*. Philadelphia: Lippincott. (Original publication).

Notsu A., O. Katai and H. Kawakami (2001). Design of Caring Communication Media Based on Nursing Theory and Interpersonal Psychology, In *the Proc. of IEEE TENCON 2001 (IEEE Region 10 International Conference on Electrical and Electronic Technology)*, pages 11–17.

Notsu A., O. Katai and H. Kawakami (2001). A Framework of Caring Interaction by a Network Model of Client's Concepts Based on a Nursing Theory and Naive Psychological Approaches, In *Proc. Of Active Media Technology 2001*, pages 279–290.

Roy, Sister Callista and H. A. Anderson (1999). *The Roy Adaptation Model, 2nd Eds.* Appleton & Lange.

Meetings and Meeting Modeling in Smart Surroundings

Anton Nijholt, Rieks op den Akker and Dirk Heylen
Centre of Telematics and Information Technology (CTIT)
University of Twente, PO Box 217
7500 AE Enschede, the Netherlands
{anijholt|infrieks|heylen}@cs.utwente.nl

Abstract

In this paper we survey our research on smart meeting rooms and its relevance for augmented reality meeting support and virtual reality generation of meetings in real-time or off-line. Intelligent real-time and off-line generation requires understanding of what is going on during a meeting. The research reported here takes place in the European 5th and 6th framework programme projects M4 (Multi-Modal Meeting Manager) and AMI (Augmented Multi-party Interaction). Both projects aim at building a smart meeting environment that is able to capture in a multimodal way the activities and discussions in a meeting room, with the aim to use this information as input to tools that allow real-time support, browsing, retrieval and summarization of meetings. In these projects many European research groups participate. Our aim is to research (semantic) representations of what takes place during meetings in order to allow generation, e.g. in virtual reality, of meeting activities (discussions, presentations, voting, etcetera). Being able to do so also allows us to look at tools that provide support during a meeting and at tools that allow those not able to be physically present during a meeting to take part in a virtual way. This may lead to situations where the differences between real meeting participants, human-controlled virtual participants and (semi-) autonomous virtual participants disappear. In this paper we introduce our research aims and ideas and we illustrate them with examples taken from many different projects in related areas.

1 Introduction

When people meet there is interaction. Interaction can be focused and it can be unfocused (Goffman 1963). Meeting means exchange of information. When two people meet this can be information about social status, by looking at clothes or posture. However, they can also start a discussion and exchange information about their family or about themselves. Whether the information exchange, or the interaction, is focused or unfocused, there need to be some common ground in order to make it effective. People meet, people gather, notice each other and communicate with each other, verbally and nonverbally, focused and unfocussed. When two people meet there usually is face-to-face interaction. Attempts to model human face-to- human face interaction have been in order to allow a translation to human-computer interaction. More recently, attempts have been made to model multi-party interaction. If more than two people meet there is another or there are others when you address a particular person. You are aware of the others and they play a role in your communication behavior, verbal and nonverbal.

There are many situations where people meet. In this paper we look at formal meetings, meetings with invited participants and with an agenda that reflects shared goals. Goals may be the willingness to discuss issues, to come to agreement and decision and willingness to accept the outcome of the meeting. Participants of such meetings see each other during different meetings, meetings where a previous meeting is summarized and discussed using its minutes. This is preferably done before starting discussions on new topics or before continuing discussions started in previous meetings. People get to know each other, sometimes know what to expect when someone takes the floor, learn about the body language of other meeting participants, learn how to interpret a participant's verbal utterances, learn about his background, his role during the meeting and

learn about his emotions and his humor. In short, meeting participants form a community. They know each other from previous meetings, they share knowledge, culture, ideas and feelings, and generally they share goals. Having shared goals allows self-disclosure during breaks, lunches or informal follow-ups of a meeting (drinks, dinners, outgoing activities, email exchanges, pictures, etc.), smoothen exchanges during next meetings.

How can we support such meeting activities? When meetings take part in smart environments, how can we make use of technology based on models of activity perception, multi-party interaction and event semantics to support meeting participants in their activities (on-line and off-line) and how can we model meeting participants as agents in such a way that remote participation or virtual participation becomes possible?

Our assumption is that people want to meet. They prefer to experience the whole gamut of activities that are associated with physical meetings and only when there are no other possibilities they seem to be willing to enter video-conferencing and computer-supported collaborative work environments. Rather than looking at ways to minimize meetings or to oblige people to use specialized meeting support technology we prefer to consider meetings as a particular case of natural interaction activity between different humans or even between humans and objects or environments. This does not mean that we don't want to distinguish between different kinds of gatherings or meetings. For example, it can be essential to know what a particular meeting is about, what the goals of the meeting or the goals of its participants are and what the reason is to have this particular meeting at this particular moment, in order to be able to understand what is going on during the meeting and, consequently, provide intelligent support to the participants of the meeting. Knowing about meeting goals helps in interpreting the actions (including the spoken utterances of the participants) during the meeting. However, it is also useful to take a more general point of view that will help us to design more advanced and attractive meeting environments.

In this paper the more general point of view is that of ambient intelligence. Ambient Intelligence has been defined as ubiquitous computing + social and intelligent interfaces. Here, 'intelligent' may refer to the original and global AI (Artificial Intelligence) paradigm, its domain-dependent specialization (as in several generations of expert systems), or its translation to agent intelligence with its distinction in believes (knowledge about an application-relevant part of the world), desires (goals of the agent in this particular part of the world) and intentions (short-term goals that bring the agent closer to its goal using a reasoning process). Interfaces between users (visitors, inhabitants) of ambient intelligence environments can be everywhere: in objects that are natural in the environment, in walls or in special devices, including PDA's or tablet PCs. Important are the social aspects of the interfaces in ambient intelligence. The environment should be able to use knowledge about our emotions, about our moods and about our personality when it tries to support us. When useful, it should be possible to induce development of social relationships between the ambient intelligence environment and its inhabitants. Moreover, usefulness of environments should not be understood in terms of efficiency or in terms of efficiency alone. Entertainment issues, feelings of enjoyment, allowing the inhabitant to feel at ease and feel comfortable are important as well.

In the next section of this paper we introduce our view on ambient intelligence and the roles of real and virtual humans in ambient intelligence environments. Section 3 is devoted to a discussion on some of the European projects in which we are involved and that have guided our insights in ambient intelligence research issues related to a virtual reality continuum. That is, we discuss how these projects contribute to the design and implementation of our view on ambient intelligence environments as discussed in section 2. We will also include observations that have become available from the Ambience project, another European project that addresses ambient intelligence issues. Section 4 is about meeting modeling. We survey our research on meeting modeling in the context of the AMI project. We zoom in on models for meeting modeling, addressee detection and the development of annotation tools. In section 5 we introduce our views (in the context of meeting situations) on the virtual reality continuum when considering meeting situations. This whole paper is an attempt to generalize from our observations obtained in the context of meetings supported by a smart environment to a context of whatever kinds of activities in ambient intelligence environments. A short summary of our findings and some notes on future research can be

found in the final section (section 5) of this paper.

2 Ambient Intelligence Requirements

As mentioned, ambient intelligence has been defined as ubiquitous computing plus social and intelligent interfaces. As may have become clear from the introduction, we are interested in the interfaces. In the ambient intelligence point of view interfaces don't need to be visible. The environment is the interface. Nevertheless, there may also be many identifiable objects and displays that can be addressed in this environment. And the inhabitant or visitor may have his or her personal assistant, available on a PDA, a tablet PC or migrating from environment to environment that can be addressed. Below are the issues we want to distinguish when looking at ambient intelligence environments.

2.1 Interpretation of Events and Activities in the Environment

This includes social and intelligent interactions in the environment between humans, between humans and objects, between humans and autonomous embodied agents (virtual humans) and interactions with the environment in general (not addressing an object or human in particular). Input can be obtained from sensors for sound, image, and haptics. The interaction that has to be perceived does not only include all aspects of focused interaction, but also aspects of unfocused interaction. Interpretation requires the fusion of all modalities that can be perceived by the environment into various levels of annotation schemes and semantic/pragmatic representations that allow further processing.

2.2 Providing Real-time Support

Based on the interpretation and the resulting representation(s) the environment, its virtual inhabitants and its smart objects need to provide real-time support to the human inhabitants or visitors of the environment. They need to decide how to present this support, through which modalities, and with which content. On the one hand there can be implicit and explicit calls for support by the inhabitant or visitor of the environment, on the other hand the environment can decide that this particular person or group of persons can benefit from its previously obtained knowledge and may suggest or perform, preferably welcome, spontaneous real-time support.

2.3 Multimedia Retrieval and Reporting

Recalling what has been going on in an ambient intelligence environment is another issue. Automatic annotation of information coming from different input sources and fusion of information coming from different input modalities into a representation that allows support to the inhabitant or visitor of an environment also allows indexing and retrieval of events, (hypermedia) browsing of activities, reporting and summarization, and a replay, e.g. in virtual reality, of what has been going on in a particular period of time or before, during and after a particularly interesting event in the environment. For the environment the collecting of such information is useful since it can help in better supporting, in real-time) its inhabitants. These inhabitants may ask such information during a gathering or the environment may supply them with this information when it considered this useful. The interests of off-line users may also guide the attention of the environment in future observations.

2.4 Autonomous and Semi-autonomous Embodied Agents

Autonomous embodied agents can be part of an ambient intelligence environment. However, we can as well have embodied agents in the environment that are real-time controlled by a distant human being or that have been sent to the environment to represent a distant human being, that is, a human not able to be there in person or to take part as a real-time controlled embodied participant of activities going on in the environment. Obviously, a human-controlled virtual being can turn into a (probably less perfect) autonomous embodied agent representing its distant owner

when it become less interesting to participate in real-time and a temporary autonomous embodied agent can change into a human-guided agent when activities require attention and real-time guidance by its distant human owner. For these applications we need to be able to present a real-time (a more or less perfect virtual reality) replay of what is happening in the environment in order to allow distant, real-time participation.

2.5 Controlling the Environment and its Inhabitants

Obviously, there can be on-line observation and participation in ambient intelligence or smart meeting environments. Capturing the events into representations that allow retrieval, browsing, summarization and multimedia generation also allows others (owners, providers, visitors) to use this information to influence and control the inhabitants and visitors of these environments. Clearly, this issue is very much related to privacy questions, that is, who has access to this information and who owns the ambient intelligence environment? The inhabitants of an environment are spied on. How does this influence their behavior? Knowing that there are eyes and ears that observe their behavior in unknown ways (details of perception, details of interpretation) may have a negative impact on natural behavior of inhabitants and visitors of ambient intelligence environments and therefore will have negative consequences for the performance of the environments. Due to these eyes and ears, available in natural objects and more or less hidden in the environment, we may even ask whether being the sole inhabitant of such an environment is in fact impossible[1]. Being there assumes to be part of a gathering and also assumes behaving as being in a public environment, including feelings of presence, co-presence, focused and unfocussed interaction behavior (Goffman 1963).

Some of these issues we discussed earlier, for instance in the context of interactive performances where human performers have to interact with objects and virtual performers in a virtual environment (see Nijholt 2000), in the context of social embodied agents (see Nijholt 2003) or in the context of presence, alienation and privacy (see Nijholt et al. 2004;Nijholt 2004). However, in particular our involvement in two European projects on meeting environments (M4 in the 5^{th} framework and AMI in the 6^{th} framework) have been fruitful in developing these ideas further, in particular the issues mentioned in the last two bullets above. For that reason we will present and discuss these projects in the next section.

3 Modeling Meetings: From Signal Processing towards Interpretation

3.1 M4: Multi-Modal Meeting Manager

In this section we first introduce the M4 project. M4 (Multi Modal Meeting Manager) is a large-scale project funded by the European Union in its 5^{th} Framework Programme[2]. M4 is concerned with the construction of a demonstration system to enable structuring, browsing and querying of an archive of automatically analyzed meetings. The archived meetings will have taken place in a room equipped with multimodal sensors.

Obviously, events and interactions that take place in a meeting room are of multimodal nature. Apart from the verbal and nonverbal interaction between participants, many events take place that are relevant for the interaction between participants and that therefore have impact on their communication content and form. For example, someone enters the meeting room, someone distributes a paper, the chairman opens or closes the meeting, ends a discussion or asks for a vote, a participants asks or is invited to present ideas on the whiteboard, a data projector presentation is given with the help of laser pointing and later discussed, someone has to leave early and the

[1] Look at remarks made by Michael Coen from MIT Labs about the effects of smart environments on their inhabitants: "The notion of being alone may disappear, or it may be changed drastically." And, "You may be in a room that's always alive and aware. And from my experiences here...when the space is 'off,' you feel it. You notice that it's not reacting. There's a void."

[2] M4 started on 1 March 2002 and has a duration of three years. It is supported by the EU IST Programme (project IST-2001-34485) and is part of CPA-2: the Cross Programme Action on Multimodal and Multisensorial Dialogue Modes.

order of the agenda is changed, etc. Participants make references in their utterances to what is happening, to presentations that have been shown, to behavior of other participants, etc. They look at each other, to the person they address, to the others, to the chairman, to their notes and to the presentation on the screen, etc. Participants have and use facial expressions, gestures and body posture that support, emphasize or contradict their opinion, etc.

The aim of the M4 project is to design a meeting manager that is able to translate the information that is captured from microphones and cameras into annotated meeting minutes that allow for high-level retrieval questions, and for summarization and browsing. In fact, but this is certainly too ambitious for the current project, it should be possible to generate everything that has been going on during a particular meeting from these annotated meeting minutes, for example, in a virtual meeting room, with virtual representations of the participants.

Figure 1: Three cameras capturing a mock-up meeting

In order to collect multimodal meeting information scripted meetings have been organized in which participants act according to prescribed rules that define periods of monologue, discussion, note taking, or a whiteboard presentation. The corpus thus obtained allows study of meeting participants' behavior. In Figure 1 we show a three-camera view of a meeting between four persons. In addition to the cameras there are lapel microphones and circular microphone arrays available for the meeting manager to capture audio. In the near future it is expected that white board pen capture can be added.

On a more detailed level the objectives of the project are the collection and annotation of a multimodal meetings database, the analysis and processing of the audio and video streams, robust conversational speech recognition, to produce a word-level description, recognition of gestures and actions, multimodal identification of intent and emotion, multimodal person identification and source localization and tracking. Models are needed for the integration of the multimodal streams in order to be able to interpret events and interactions. These models include statistical models to integrate asynchronous multiple streams and semantic representation formalisms that allow reasoning and cross-modal reference resolution. These models form the basis of browsing, retrieval, extraction and summarization methods. Textual "side information" (the agenda, discussion papers, slides) enables the application of useful constraints. It may be used to adapt the language model of the speech recognizer or as query expansion information for retrieval.

A straightforward meeting browser can follow the structure of an agenda. Each agenda item can be associated with different views on that topic. For example, a textual summary, a diagrammatic discussion flow indicating which participants were involved (speaker turn patterns), and audio and video key frames that give the essence of the discussion. Obviously, in order to track the discussion and find the interesting parts features need to be distinguished that can be recognized by the meeting manager.

Presently there are two approaches that are followed. The first one is the recognition of joint behavior, that is, the recognition of group actions during the meeting. Examples of group actions are presentations, discussions, consensus and note taking. Probabilistic methods based on Hidden Markov Models (HMMs) are used for this purpose (McCowan et al. 2003). The second approach is the recognition of the actions of the individuals independently, and to fuse them at a higher level for further recognition and interpretation of the interactions. When looking at the actions of the individuals during a meeting several useful pieces of information can be collected. First of all, there can be person identification using face recognition. Current speaker recognition using multimodal

information (e.g., speech and gestures) and speaker tracking (e.g., while the speaker rises from his chair and walks to the whiteboard) are similar issues. Other, more detailed but nevertheless relevant meeting acts can be distinguished. In Zobl et al. 2003 recognition of individual meeting actions by video sequence processing in the context of the M4 project is discussed. Examples of actions that are distinguished are entering, leaving, rising, sitting, shaking head, nodding, voting (raising hand) and pointing (see Figure 2). These are rather simple actions and clearly they need to be given an interpretation in the context of the meeting. Or rather, these actions need to be interpreted as part of other actions and verbal and nonverbal interactions between participants. Presently models, annotation tools and mark-up languages are being developed in the project that allow the description of the relevant issues during a meeting, including temporal aspects and including some low-level fusion of media streams. Higher-level fusion, where also semantic modeling of verbal and nonverbal utterances is taken into account has not been done yet. In some cases it turns out to be more convenient to make shortcuts to a pragmatic level of fusion using knowledge from the application.

Figure 2: Pointing, rising and voting

The M4 meeting manager captures the events and interactions in the meeting room. After capturing the gathered information becomes off-line available for both participants and non-participants. Clearly, we can look at the project as research on smart environments and on ambient intelligence. However, there is no explicit or active communication between user and environment. The user does not explicitly address the environment, although it would be possible, but not done in this project, that a meeting participant explicitly addresses the meeting manager the way she would address a human note taker during a meeting. Currently, the environment registers and interprets what's going on, but is not actively involved. The environment is attentive, but does not give feedback or is pro-active with respect of the users of the environment. Real-time participation of the environment requires not only attention and interpretation, but also intelligent feedback and pro-active behavior of the environment. It requires also presentation by the environment of multimedia information to the occupants of the environment.

Our involvement in the project is modest and it should be understood that most of what we explained above is based on work done by our partners. In our work, see e.g. Jovanovic (2003), we try to explore different aspects of the interpretation point of view. In addition we hope to integrate recent research in the area of more traditional multimodal dialogue modeling (Hofs et al. 2003). These issues will become more important in the recently started AMI project, an overlapping successor project of M4.

3.2 AMI: Augmented Multi-party Interaction

The AMI (Augmented Multi-party Interaction)[3] project is concerned with new multimodal technologies to support human interaction, in the context of smart meeting rooms and remote meeting assistants. The project aims to enhance the value of multimodal meeting recordings and to make human interaction more effective in real time. These goals are being achieved by developing new

[3]AMI started on 1 January 2004 and has a duration of three years. It is supported by the EU 6^{th} FP IST Programme.

tools for computer supported cooperative work and by designing new ways to search and browse meetings as part of an integrated multimodal group communication, captured from a wide range of devices. The project also makes recorded and annotated multimodal meeting data widely available for the European research community, thereby contributing to the research infrastructure in the field.

In the next paragraphs we introduce the AMI project. Clearly, since the project has to start yet, we have to confine ourselves to the project proposal and the different research tracks that have been defined there. From the point of view of the virtual reality continuum (see the next section) the following tracks are especially relevant:

Understanding Meetings: Which meeting characteristics play a role in order to understand the group's communication? Multimodal turntaking dynamics and multi-party interaction modeling are general areas of research. How do turntaking and dialogue structure depend on these meeting characteristics? Examples of characteristics are size, status differences, familiarity with each other, the setting, the goal or task (maintaining sociality, sharing information, generating ideas), etc. Although presently M4 is about face-to-face discussions, other meeting modes, supported by communication technology, can be considered, for example allowing asynchronous communication or video-conferencing. That is, in AMI not only face-to-face but also remote meeting dynamics has to be studied. Clearly, a wealth of research has been done in these areas and can be made use of, but in addition to that meeting support research, here we need also the environment to understand the meeting in order to allow later access for retrieval, replay and explanation.

Uni- and Multi-modal Recognition: There are many challenges for audio and video processing in smart environments. There are multiple sound sources, speech is conversational and there may be non-native speakers, to mention a few problems for speech recognition. For video processing we have to deal with unrestricted behavior of participants with variations of appearance and pose, different room conditions, occlusion, etc. Speaker turn detection, speaker localization and speaker tracking can be done using speech recognition and identification; visual processing is needed for visual tracking, face detection and recognition, facial expression recognition, gesture and action recognition. However, multi-channel processing, i.e., combination of audio and video streams allow better and more complete person identification and tracking and understanding of human-human interaction in a smart meeting environment. Multimodal syntactic and semantic information need to be extracted in order to recognize and interpret participant behavior, participant interaction and meeting events.

Multimodal Content Abstraction and Multimedia Presentation: Retrieval from meetings and browsing of meetings requires a natural structuring of meeting content. This structuring is obtained from recognition and interpretation of sequences of meeting acts and indexing the multimodal recordings. Some example questions that the AMI demonstration system should be able to answer are: Who were the participants? Was the agenda covered? How did the discussion progress? What was the atmosphere? Can I have a summary of the meeting? Segmentation of a meeting can be done from different viewpoints. We can look at events such as discussion, monologue, note taking, presentation (as is already done in the M4 project), but also at a structuring in terms of decision points, task assignments and topic shifts. An intelligent meeting browser can be designed that uses a hypertext view of the meeting in which the different structuring viewpoints are embedded.

Remote meeting assistant: One of the issues that will be explored in the AMI project is the design of a real-time, on-line remote meeting assistant. The system will allow a remote participant to a meeting to browse recent events in the meeting or to be automatically alerted at points of interest. Obviously, this empowerment of a remote participant can be useful for others present at the meeting too.

3.3 Related Research Projects

There have been several other research projects concerned with the computational modeling of meetings or, more modestly, the development of tools that help to support meetings or to off-line review and retrieve information available in recordings of meetings. For example, the ICSI project is also concerned with the development of a system for recording and browsing meetings, however, it is only based only on audio data (Morgan et al. 2001). A project very much related to M4 is the Meeting Room project at Carnegie Mellon University (Schultz et al. 2001). It is concerned with the recording and browsing of meetings using audio and video data. Closely related to AMI is for example the work done at the University of California, San Diego, which includes the development of methods for person identification, current speaker recognition, models for face orientation, semantic activity processing and graphical summarization of events. There is both work on intelligent meeting rooms (Mikic et al. 2000) as on smart environments in general (AVIARY: Audio-Video Interactive Appliances, Rooms and sYstems see Trivedi et al. 2000). Neem (Ellis and Barthelmess 2003) is a project of the University of Colorado that aims at introducing different intelligent agents in a distributed business meeting environment. These agents have to assist the meeting participants. Three agents are considered: an informing agent (assisting in obtaining necessary information, e.g. through a web search), a social agent (helps to build common ground) and an organizational agent (keeping track of time, etc.). Underlying their behavior is Bales' Social interaction Systems (Bales 2001) theory and organizational theories of problem solving. The Ambiance project, done in the context of a European project, is also more general than 'just' an attempt to model meeting situations. Rather it looks at smart home environments (Aarts et al. 2003), requiring much more modeling of the environment, including the many objects that can play a role in activities among inhabitants or between inhabitants and the global environment.

4 Meeting Modeling

In this section we have a few preliminary observations on meeting modeling. The various behaviors of peoples in a meeting can be analysed and studied from different perspectives. Meetings are social events: familiarity, social roles, personalities influence the behavior of participants. In many meetings a group meets to work on a project, conversations take place that have the form of a discussion. The task of the group implies taking decisions what to do to reach the goals of the project, and often to become clear about the goals of the project. Thus, an important part of a meeting model, a model that describes the joint meeting activities, is a discussion model. We could look at the meeting as just a series of conversational, verbal or non-verbal behaviors, observe for instance turn taking and turn giving behavior, or see how topic change is realized, or how participants address other participants, but we feel that without taking into account the goals that the participants want to realize by meeting we can not fully understand their behaviors and the joint activities that take place. It is the goal of the group and the -possibly conflicting- interests of the participants that finally motivate what is being said and how people react on each other.

To give a concrete example, consider the following situation. After a student has given his final presentation of his master thesis, a small group of people, involved in the student's project, has to judge the student's work. The judgment has to be expressed in the form of a mark on a scale between 5 and 10. The four people meet and they have about a quarter of an hour to come to a decision; the student is waiting outside the meeting room for the outcome of the decision.

Although this is a rather simple situation: the topic of discussion is clear, the possible outcomes of the process are clear and fixed in advance, and there may even be prescriptions what aspects have to be taken into account for such a decision, many of the ingredients of discussions in which a group has to make a decision can be observed in this situation. The following question may be of interest for such a proces.

- How was the decision made?
- Did all members agree on the outcome?

- How long did it take before a decision was made?
- Was the discussion well organized and structured or were there many topical shifts?
- Did everyone have the chance to give his opinion?
- Was there interaction between participants having different opinions?
- Were there argument given for or against statements?
- Was there a discussion about the criteria that had to be taken into account?
- Was there a discussion about the weights of the different factors that were of influence on the outcome?
- Was there a group member who was convinced by other members and changed his opinion?
- Was every member evenly involved in the discussion or were there clearly distinguished parts in which some members showed more involvement than others?

Notice that we don't ask whether the outcome of the decision making process was a *rational* one.

Of interest is the way the group comes to an agreement, not whether the conclusion is a reasonable or logical conclusion. The model is a *descriptive* model not a model that prescribes how the participants *should* behave or discuss, or how they should come to a conclusion.

If we observe all relevant information, and aspects of conversational behavior, of a large number of similar groups making the same task, we can compare the results and see what factors influence the outcome of the decision and the time it took the group to come to a decision.

A meeting model should be general enough so that it models not only one type of meeting in which a group discusses one specific topic, but whatever topics and issues that are discussed. The meeting model needs a model of discussions in general. What are the basic elements of a discussion and how are they structured?

A discussion has a topic: the issue the discussion is about. The topical structure of a conversation show where a subtopic or a new topic is introduced and by whom. The discussion starts when someone gives his opinion, explains his position and gives the floor to other participants to give their opinion about it. We can distinguish a number of types of contributions to the discussion: give a new statement or opinion, react on a previous given statement, either by agreeing or disagreeing, or by partial agreeing with the statement. One can ask someone for his opinion, or ask for clarification. Finally, one can ask whether every one agrees on a particular conclusion.

For all these types of actions people use verbal and non-verbal expressions to communicate them.

The higher level information stored in the model, that is the information on the level of the discussion and the decision making process, is 'backed up' by information about events on a lower level of general speech acts and conversational behavior: the transcripted speech, the voice and prosody of the speech, the information about nonverbal conversational behavior, like head nodding, pointing gestures. From observations obtained form data received over the video and audio channels we may conclude for instance that the speaker is person A and that he strongly disagrees with the current statement of the discussion.

We may view the meeting browser as an interface of an meeting expert system that can be asked not only to give information about an event but also to show the audio and video data that together form the evidence for its conclusions about what happened in terms of semantic actions.

The *back up* relation between the information on the higher semantic level of group processes and the information on the lower level of individual behaviors that take place is one of the types of relations that exists between the various actions that we distinguish in the meeting model. The *constitutive* relation between two types of actions x and y is that relation that we express by saying that a person is doing x by doing y. An example is: switch on the light by pressing a button. Another is to vote by raising the arm. The constitutive relation can either be conventional (ritual)

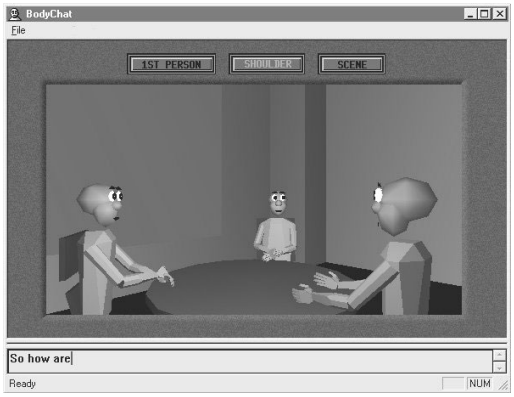


Figure 3: BodyChat: Conversational gestures

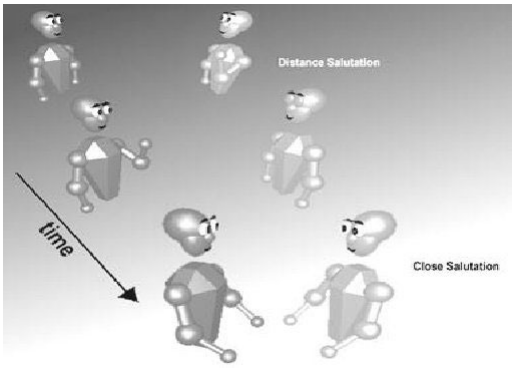


Figure 4: People meet: Salutations

or natural and based on causal relations between the two events. Other types of relations between actions are the *sequential* relation; action: x is followed by action y, and the *joint-relation*: x and y are simultaneous actions on the same level that together constitute one joint-action: shaking hands is a typical example.

5 Meetings in a Virtual Reality Continuum

As may have become clear in the first sections of this paper, developments in the area of ambient intelligence or in more restricted environments such as smart meeting rooms and future workspaces have drawn attention to the modeling of multiparty interaction, where the members of the party may be human only or, when smart objects and other support technology become available, both humans and objects. There is an obvious trend in meeting support technology to allow remote participants or to only have geographically distributed meeting participants. This has been the start of research on video conferencing and collaborative environments where attempts were made to provide information about gaze in order to facilitate the turntaking process (see e.g. Vertegaal 1998). Again, in ambient intelligence environments and certainly in smart meeting rooms similar research issues emerge with the aim to understand behavior, interactions and events, while making use of audio, video and biometric sources. As mentioned before, this information may as well be used to generate virtual reality representations of meeting participants in a virtual meeting room or an augmented reality supported physical meeting room. Meeting participants can be physically present, they can be represented by an (embodied) agent that alerts and supports when things become interesting (just as the remote meeting assistant) - but otherwise is rather passive - or they can be immersed in the (distributed) virtual environment together with the other participants, all represented as avatars mimicking their owners.

In the subsections below we show some examples from the literature and some of our own research.

5.1 Multi-party Interaction: BodyChat and Situated Chat

hesubfigure In virtual reality environments examples of research on multi-party interaction can be found. For example, Vilhjálmsson (1998) has worked on BodyChat (Figure 3, Figure 4), a chat environment system that allows users to communicate via keyboard input, "while their avatars automatically animate attention, salutations, turn taking, back-channel feedback and facial expression, as well as simple body functions as the blinking of the eyes." Hence, human-like conversational behavior for virtual humans that represent real users is simulated. In this system, apart from what is derived from the situation and the utterances, there is not necessarily a relationship between what a particular chat participant is doing in real-life (posture, gestures,

facial expressions) and its nonverbal communication characteristics in the virtual world. It is the avatar that knows how to use his body during communication. This work has been continued in a project called Situated Chat (Vilhjálmsson). In addition to the social conversational rules Situated Chat also used a discourse context model to automatically generate referring gestures in the shared visual environment of the animated avatars.

Translation of this work to a smart meeting environment is straightforward. Once we can capture the events in a physical meeting room we can translate them to events in a virtual meeting room (see e.g. Figure 5) and add remote participants or add model-based behavior to virtually represented participants. For example, focus tracking (Stiefelhagen 2002) can be enhanced and converted into gaze behavior of virtual meeting participants. Assigning desirable properties to avatars that represent human participants during a meeting may much more smoothen the progress of a meeting than when the real participants are represented with all their particularities. This view allows a particular participant to become more lively through more extrovert gestures and facial expressions, it allows to convert a non-native speaker to a native speaker and it even allows to change the physical appearance of a particular participant.

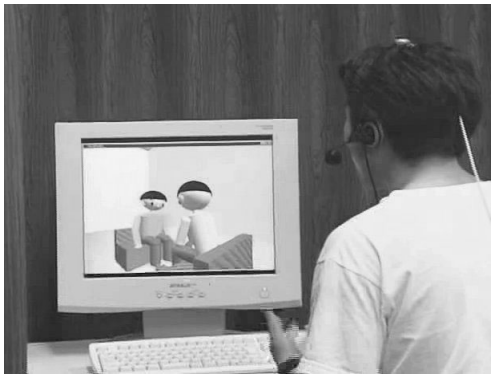


Figure 5: Real-time transformation of conversational gestures

6 Multi-party Interaction: Mission Rehearsal Exercise

Another example, where the starting point is the virtual world inhabited by autonomous agents, is the Mission Rehearsal Example (MRE) environment (Traum and Rickel 2001) developed at the Institute for Creative Technologies.

Figure 6: Multi-party interaction in the Mission Rehearsal Exercise

This training environment allows immersive participation in multi-party interaction. In this system there are autonomous agents in a virtual world that are able to interact with a human visitor (in this case, a trainee that has to perform a certain task) that is immersed in the environment. There is direct interaction (the trainee addresses a particular agent he sees in the environment) and indirect interaction (the embodied agents in the environment have their own tasks, not everybody is always involved in every interaction). See Figure 6. Hence, we have multimodal interaction between multiple (human and virtual) agents in the environment. Important are the locations of the conversants and the objects they are discussing. Agents are aware that others are listening. An important aspect of this system is the underlying dialogue model. It consists of several layers: a contact layer (whether and how individuals are accessible for communication), an attention layer (the objects or process that agents attend to), the conversation layer (where separate dialogue episodes are modeled), a layer of social commitments and a layer of negotiation (how agents come to agree on commitments). Although the models are there it is certainly not yet the case that in this environment there is free interaction between the multiple (virtual and human) agents. Currently the layered model underlies a scripted interaction.

A similar environment for learning Lebanese Arabic language and culture is being developed

Figure 7: Tactical Language Training project

Figure 8: Virtual presenter

at the CARTE institute. The environment is inhabited by animated agents representing local people with who a learner has to communicate (see Figure 7). The learner is also represented in the environment where his avatar displays the chosen gestures.

6.1 The HMI-Parlevink Virtual Meeting Environment

The AMI project just started. On the other hand, our research group has a background in modeling embodied agents in their 'natural' environments. Some preliminary research on modeling meeting behavior and displaying it in a virtual meeting room is under way. One of the topics we look at this moment is the role of a virtual presenter in a virtual meeting room. Previous work in this area has been done by Nomay et al. 2000 (see Figure 8). One thing we would like to model is to have a remote participant showing a presentation as an embodied agent. It certainly should allow interaction with this embodied representation and probably also with the remote participant who is (semi-)controlling the presentation. However, in our situation we also want to allow fully synthetic presenters that know about the presentation and that are able to interact with meeting participants (maybe present in a physical meeting room, maybe a remote human meeting participant and maybe a fully synthetic virtual assistant). Can we interrupt this synthetic presenter while he or she is showing a PowerPoint presentation? The presenter knows about all sheets in the presentation and should at least be able to tell that the answer to a particular question will be on a next sheet. Or that he or she has already handled that, but is willing to go back to a particular sheet in order to explain it again or in more detail.

Apart from increasing the notion of (real-time) presence, when we combine virtual, real, mixed and augmented meeting settings, there is also the notion of validation of theories of meeting interactions by looking at possibilities to generate such interaction behavior from models of interaction or from (semi-)automatically obtained annotations from meeting interactions.

6.2 Putting it All Together

In this section we made clear that some modest research attempts are underway to achieve models that cover verbal and nonverbal communication aspects of human behavior in different situations. These models are necessary to allow for a smooth transition from real to virtual worlds and to a merging from real and virtual worlds. Due to our participation in European projects on meeting modeling, meeting situations and meeting interactions our main efforts are in the area of meetings. However, there are so many different kinds of meetings, meeting situations, meeting interactions and meeting participants that we don't think this domain very much restricts our interest in modeling human interaction in all possible kinds of situations.

7 Conclusions

We discussed different application areas where it has become useful to model multi-party human interaction behavior. Our main observation in this paper is that we see research in previously separate areas converge and that there is a natural trend towards situations where ambient intelligence environments (exemplified in this paper with smart meeting rooms) and virtual reality environments merge in order to obtain shared environments where people live, work and meet. In this paper we surveyed our research and research ideas in the framework of the European AMI (Augmented Multi-party Interaction project). We hardly touched upon our technical work in this project.

Figure 9: Meeting audience showing appreciation

Apart from meeting modeling (see section 4) we are in particular concerned with the design of annotation tools, image processing (posture, gesture and facial expressions), modeling of turntaking and addressee detection and emotion modeling, all in the context of meetings in smart environments. There is a lot of research that is extremely important, but is not discussed here and not part of the project. We would like to mention privacy issues, presence issues and issues related to the fact that people know that there actions are recorded and interpreted (cf. Nijholt 2004). Presence issues in a meeting environment have been researched by Slater (Pertaub et al). Slater studied the illusion of sentience in a virtual meeting environment with the objective to present evidence that people react to virtual characters as if they were real. See Figure 9 where someone is presenting for an interested (virtual) audience. Obviously, these observations are interesting when we allow mixtures of virtual and (representations of) real people in the same meeting environment.

References

Aarts, E., R. Collier, E. van Loenen & B. de Ruyter (Eds.). Ambient Intelligence. Proceedings First European Symposium, EUSAI 2003, Lecture Notes in Computer Science, Springer, Berlin, 2003.

Bales, R.F. *Social Interaction Systems. Theory and Measurement.* Tranaction Publishers, New Brunswick, 2001.

Ellis, C. and Barthelmess. The Neem dream. Proceedings *Tapia '03*, October 2003, Atlanta, Georgia, USA, 23-29.

Goffman, E. *Behavior in Public Spaces.* Notes on the Social Organization of Gatherings. The Free Press, New York, 1963.

Hofs, D., R. op den Akker & A. Nijholt. A generic architecture and dialogue model for multimodal interaction. Proc. 1st Nordic Symposium on Multimodal Communication, P. Paggio, K. Jokinen & A. Jönsson (Eds.), CST Publication, Center for Sprokteknologi, Copenhagen, 2003, 79-92.

Jovanovic, N.. Recognition of meeting actions using information obtained from different modalities: a semantic approach. TR-CTIT-03-48, October 2003, 44 pp.

McCowan. I, S. Bengio, D. Gatica-Perez, G. Lathoud, F. Monay, D. Moore, P. Wellner and H. Bourlard. Modeling Human Interaction in Meetings. Proc. *IEEE ICASSP 200*3, Hong Kong.

Mikic, I., K. Huang & M. Trivedi. Activity monitoring and summarization for an intelligent meeting room. In: Proceedings IEEE Workshop on Human Motion, Austin, Texas, December 2000.

Morgan, N., D. Baron, J. Edwards, D. Ellis, D. Gelbart, A. Janin, T. Pfau, E. Shriberg & A. Stolcke. The Meeting Project at ICSI. Human Language Technologies Conference, San Diego, March 2001.

Nijholt, A. Towards virtual communities on the Web: Actors and audience. Proc. *Intelligent Systems & Applications (ISA'2000)*, Vol. II, F. Naghdy et al. (Eds.), ICSC Academic Press, Canada, 2000, 725-731.

Nijholt, A. Multimodality and Ambient Intelligence. In: *Algorithms in Ambient Intelligence*. W.F.J. Verhaegh, E.H.L. Aarts & J. Korst (Eds.), Kluwer Academic Publishers, Boston/-Dordrecht/London, 2003, 21-53.

Nijholt, A., T. Rist & K. Tuinenbreijer. Lost in ambient intelligence? In: Proceedings *ACM Conference on Computer Human Interaction (CHI 2004)*, April 2004, Vienna, Austria, ACM New York, 1725-1726.

Nijholt, A. Where computers disappear, virtual humans appear. *Computers and Graphics*, Vol. 28, No. 4, Elsevier, ISSN 0097-8493, 2004, to appear.

Nomay, Ts., L. Zhaoz & N.I. Badler. Design of a Virtual Human Presenter. Internal report, University of Pennsylvania, 2000.

Pertaub D.-P, M. Slater & C. Barker. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments* 11 (1), 68-78.

Schultz, T., A. Waibel, M. Bett, F. Metze, Y. Pan, K. Ries, T. Schaaf, H. Soltau, M. Westphal, Hua Yu & K. Zechner. The ISL Meeting Room System. Proceedings of the Workshop on Hands-Free Speech Communication (HSC-2001), Kyoto Japan, April 2001.

Stiefelhagen, R. Tracking focus of attention in meetings. Proc. IEEE *International Conference on Multimodal Interfaces*, Pittsburgh, PA, USA, 2002, 273-280.

Traum, D. and J. Rickel. Embodied agents for multi-party dialogue in immersive virtual worlds. In *Agents 2001 Workshop on Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents*.

Trivedi, M., I. Mikic, S. Bhonsle. Active Camera Networks and Semantic Event Databases for Intelligent Environments. IEEE Workshop on Human Modeling, Analysis and Synthesis (in conjunction with CVPR), Hilton Head, South Carolina, June 2000.

Vertegaal, R. Look who's talking to whom. Mediating joint attention in multiparty communication & collaboration. Ph.D. Thesis, University of Twente, 1998.

Vilhjálmsson H., and J. Cassell. BodyChat: Autonomous Communicative Behaviors in Avatars. In: Proc. *2nd Annual ACM International Conference on Autonomous Agents*, Minneapolis, 1998.

Vilhjálmsson, H. Avatar Augmented Online Conversation, Ph.D. dissertation, Program in Media Arts and Sciences, Massachusetts Institute of Technology, Cambridge, MA.

Zobl, M., F. Wallhoff & G. Rigoll. Action recognition in meeting scenarios using global motion features. Proc. IEEE International Workshop on *Performance Evaluation of Tracking and Surveillanc*e, 2003.

A Research Environment for Meeting Behavior*

Wilfried M. Post, Anita H.M. Cremers and Olivier Blanson Henkemans
TNO Human Factors
P.O. Box 23, 3769 ZG
Soesterberg, The Netherlands
{Post|Cremers|Blanson}@tm.tno.nl

Abstract

To understand meeting behavior better and to make meetings more useful and satisfactory, we are developing an experimental environment for studying meetings in a controlled manner. In this environment, we will be able to control and manipulate the various factors that influence efficiency and participant' satisfaction. The meeting behavior is measured in terms of both the process during the meetings and the outcomes. The research environment is based on an existing framework for studying the impact of technology on groups. We specify the research environment and provide an example scenario. We discuss two main applications of the environment. The first one is explorative, in the context of application development. We are especially interested in developing and evaluating innovative, ICT supported ways of conferencing. The second one is experimental, in the context of theory development.

Keywords: Meetings, Natural Interaction, Design, User interface design

1 Introduction

In most organizations, meetings are a generally accepted mean for information exchange and decision-making. Although meetings are beneficial for common achievement of goals, they are also often considered not to lead to clear-cut results and to be time-consuming and unrewarding (Piët (1990)). This may be caused by several factors, such as incompetent chairperson, ill-informed participants or lack of an agenda. To make meetings more effective and efficient as well as more fruitful and satisfactory for participants, factors should be determined that influence meeting behavior. This may lead to new meeting concepts and IT support tools. Subsequently, the effects of these on actual meeting behavior should be studied.

In order to be able to study meeting behavior in a controlled manner, we are developing a research environment in which we can evoke and manipulate particular kinds of meeting behavior and can monitor and measure the meeting process and outcomes. This enables us to examine the effects of specific factors, ranging from organizational aspects to the use of particular meeting tools. Our ultimate goal is to improve our understanding of meeting behavior and to develop models and theories for it.

In this paper, we first present our ideas on meeting issues in the form of a meeting paradigm. Next, we describe the research environment and give and example of how it could be used to investigate meeting behavior. Finally, we provide an overview of research issues that can be addressed using this paradigm. We conclude by discussing the advantages and drawbacks of our approach.

2 Meeting Paradigm

What is our scope in the study of meeting behavior? Meetings are regarded as a means to reach a higher-level goal, such as the coordination of activities, the development of a product, or solving

*This work was partly supported by the European Union 6^{th} FWP IST Integrated Project AMI (Augmented Multi-party Interaction, FP6-506811, publication AMI-24).

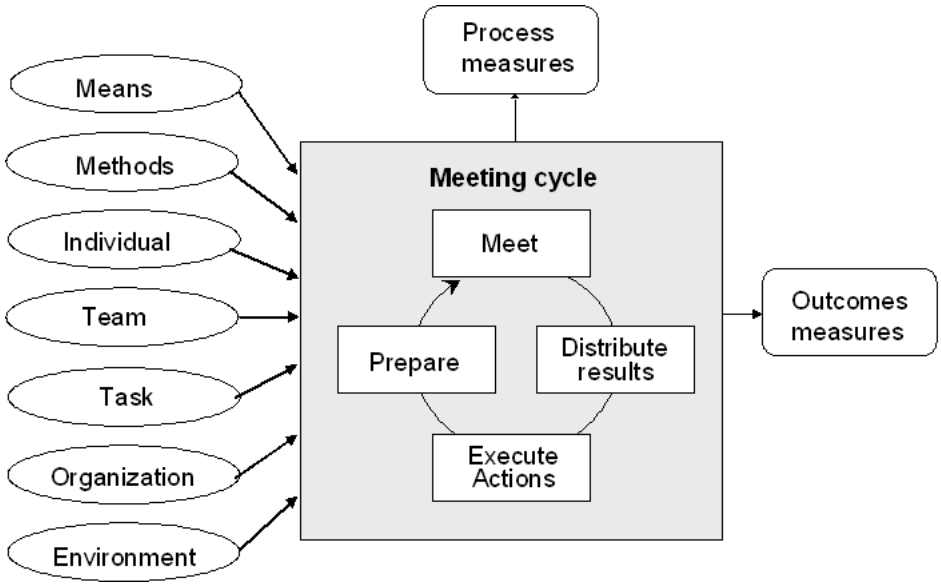


Figure 1: Conceptualization of the meeting paradigm. The process and the outcome of cycles of meetings depend on several input factors.

a crisis. Meetings should not be considered as isolated events. Meetings are usually followed by distribution of the results, execution of actions that have been agreed on and preparation of the next meeting. This meeting cycle can be infinite, such as the meetings of a management board, but also finite, such as within a project or an operation. We propose a meeting paradigm that is based on this meeting cycle, and on aspects we have identified from the literature. We have organized these into seven basic factors. The categorization is based on the conceptual framework for studying the impact of technology on groups, as proposed by Short et al. (1976). Fig. 1 conceptualizes our meeting paradigm. This conceptualization should not be seen as a model of meetings, but as a mechanism for generating and measuring meeting behavior. It allows us to manipulate the various input variables and to measure the process and the outcome of the meeting cycle. Process measures are taken during the execution of a cycle at particular points in time (e.g., directly after a meeting). For instance, the status of the individual and shared information and knowledge, workload, group cohesion, etc. may be determined with measuring instruments such as questionnaires and tests. Outcome measures may include the speed and timeliness, the quality and quantity of the end product, and the satisfaction of the participants.

We distinguish *means, methods, individual factors, team factors, task factors, organizational factors and environmental factors*. *Means* refer to systems and tools that support a meeting-related tasks (e.g., an interactive large screen display). *Methods* refer to prescriptions of *how* to do a particular task (e.g., a procedure to chair a meeting). A *team* is a group of individuals who see themselves and are seen by others as a social entity (Guzzo and Salas (1995)), which is also the case for the participants of a meeting (e.g., a management team). Team processes are influenced by *individual* characteristics (Shaw (1971)), in particular the different roles that the *individuals* play (Stewart et al. (1999)) (e.g., the chair). The *task* refers to the work that must be done to reach certain goals. Through the task, team members become interdependent (Steiner (1972)). Tasks can be described as individual tasks (e.g., design) (Steiner (1972)) or as group task (Short et al. (1976)) (e.g., negotiate). *Organizational factors* refer to aspects such as organizational structure and culture (Swigger et al. (2004)). *Environmental factors* refer to aspects external to the organization, such as the market. The success of teams and organizations strongly depends on how they manage the unexpected dynamics of the environment (Weick and Sutcliffe (2001)).

3 Experimental environment

In order to examine meeting behavior and outcomes in the meeting paradigm, the environment should generate *replicable* meeting behavior and at the same time facilitate *natural* meeting behavior. The approach that we follow is to define *the provisions for the emergence of* meeting behavior. For each application of the environment, all input factors should be specified: what means and procedures are used, what roles the individuals play, the configuration of the team, a definition of the task, the specifics of the organization, and the environmental developments. Participants are asked to carry out a certain task (i.e., to reach a common goal), and we make them interdependent by providing role-restricted information. This requires that they all need to confer about the same matter, but it does not prescribe how and when.

As is the case in many organizations, the subjects can use e-mail, a web browser, and PowerPoint for information presentation. However, we maintain control of the way they can exchange information. We do this by simulating an office environment, in which we control the tools they use, the information they can find on a simulated web, and simulate e-mail communication from individuals in the environment (such as head of departments, and parties outside the organizations). We will illustrate this approach below in the form of a scenario focusing on design project meetings.

3.1 Example Scenario

The example scenario focuses on design project meetings. These types of meetings are common in many (commercial) organizations that will certainly benefit from a more efficient process and a higher quality outcome. Due to our design research (Punte and Hamel (2001)) and our involvement in design projects, we are familiar with the design task and know what real-life design behavior should look like. Further, since a project is finite, and a concrete product is designed, measuring the process and outcome of meetings is expected to be relatively straightforward.

The proposed scenario is instantiated as follows: a dynamic market (changing fashion), an organizational context (a particular company, budget cuts), a task (a design project), a team (consisting of 4 participants), individual characteristics (with particular roles), a method (four defined design phases), and means (laptops with the e-mail, a web browser, PowerPoint, four individual offices and a meeting room including a beamer). The details of the scenario, which is illustrated in Fig. 2, are described next.

Four subjects, acting as employees of a consumer electronics company, join a project on designing an innovative TV remote control (RC). The project roles are: project manager (PM), marketing expert (MA), user interface designer (UID) and industrial designer (ID). The overall project method that has to be followed has three phases: functional design, conceptual design, and detailed design. Each phase is followed by a meeting.

After having prepared their role individually (for which they receive a short training) and the first meeting, they meet, bringing their role-specific laptops. The participants get acquainted, and the project manager starts the meeting officially by providing the project plan and the division of work until it is clear to everyone. After the meeting, individual work is carried out, including the preparation of the next meeting.

During the preparation of the second meeting, PM gets e-mails on budget cuts, MA receives a marketing report with user requirements and desires, UID devises the remote control functions, based on examples found on the (simulated) web, and the ID devises the functionalities of the RC, also inspired by the web. They all prepare (pre-structured) PowerPoint presentations. During the next meeting they exchange their findings and ideas, and come to an agreement on the functional design.

They then split up again, to carry out individual work. Now PM gets e-mails about deadline changes, MA gathers market changes and evaluation criteria on the web (fruit is the fashion, yellow the most popular color), UID finds examples of old and new RC interfaces (such as scroll wheels and speech control), and the ID on components, properties and materials. During the meeting that follows, they present their PowerPoint slides, and try to reach agreement on the conceptual design, also dealing with the changing project constraints and market.

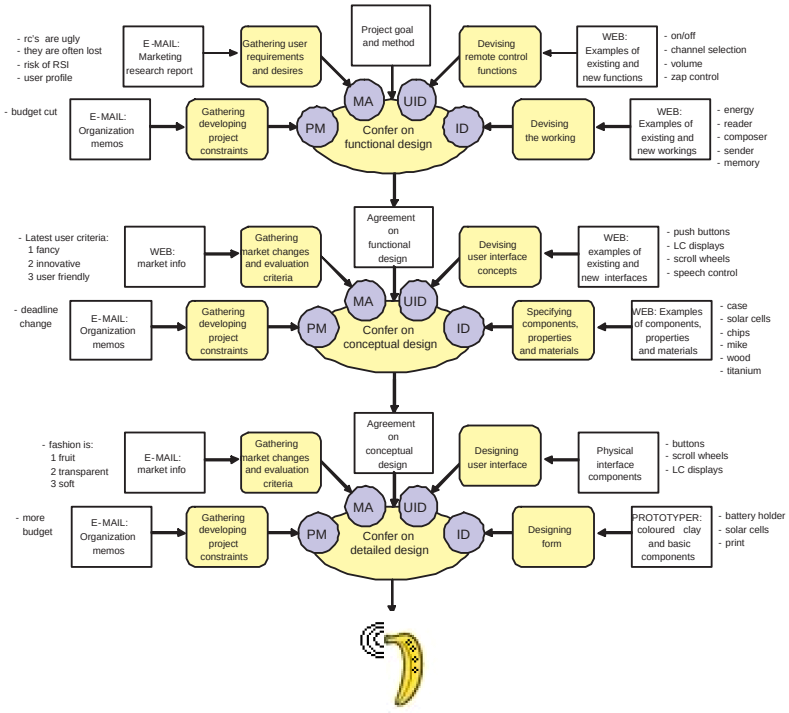


Figure 2: A scenario for designing an innovative TV remote control. Square boxes denote information, rounded boxes individual activities, ovals group activities, circles roles. PM: project manager; MA: marketing expert; UID: user interface designer; ID: industrial designer.

The last phase starts with individual work again. PM gets more budget, MA develops an evaluation scheme, and UID and ID work together on a clay prototype. They present their prototype called "Chiquita", which is assessed according to the criteria of MA. The project is concluded with a small party.

During the project, various process measures can be obtained. Information use can be logged (i.e., opening an e-mail or a web page). At certain points in time, subjects can be asked automatically to fill out questionnaires, e.g. work load. At the end, the outcome of the project can be measured. Since the scenario will contain objective criteria for the design, the quality of the design can be determined. We can also measure the speed and timeliness of the project, and the satisfaction of the users. Finally, depending on the research question, we can observe and code specific behavior of the participants.

4 Application of the Paradigm

The above example shows how meeting behavior can be generated in a controlled but not in a pre-scribed manner and how process and outcome measurements can be collected. This semi-natural setting can be used in two ways: exploratory as well as experimentally.

4.1 Explorative setting

In an explorative setting, variants of input factors at all levels can be introduced, to monitor and measure the effects of these factors. Based on process and outcome measures, the factor variant can be adjusted and introduced in a new meeting cycle in an iterative manner, to ultimately come to the variant that optimally serves the goals of the project team. An example of an explorative setting is the development of certain means to support the participants during the meeting cycle, such as a meeting browser. A meeting browser is an application that supports users in finding

elements of interest in multi-media digital recordings that have been captured during previous meetings (for example Chiu et al. (2001);Guillemot et al. (2003)).

The paradigm could help designing the meeting browser in an iterative manner, similar to the design process that has been described in the scenario. At an early stage of the development process, a meeting cycle could be organized to elicit user requirements for the meeting browser. In later phases of the design process, functional, conceptual and detailed designs of the meeting browser could be introduced consecutively in the sessions. Outcome and process measures could be used to improve the design to ultimately result in a meeting browser that optimizes the process and outcome of project meetings.

In every phase, specific methods could be applied to measure effectiveness, efficiency and satisfaction of users (i.e. meeting participants) with respect to the current version of the meeting browser. In the initial phase, the paradigm provides a framework for determining the functions of the meeting support tools to be developed. It can help establishing characteristics of future users, their information needs and the ways they would like to search for this information. Different measurements can be applied to elicit requirements and evaluate early system designs. In contrast to methods for user requirement elicitation for meeting browsers that have been applied earlier (e.g. Lisowska (2003)), the proposed method provides the possibility to, non-obtrusively, acquire requirements in a semi-natural setting. Before a meeting cycle starts the characteristics and background knowledge of participants can be determined by means of a questionnaire. During meetings researchers may observe the process and record naturally occurring information needs of participants and the ways in which these are expressed. Notes that are made by participants during meetings form a valuable source of information to determine what they found important and maybe were afraid to forget. In order to distribute the results, meeting participants can be asked to give an account of the meeting to a colleague who was not present. Again, this is a valuable source of information to know what participants remember of a meeting, how they remember it, and even more important, what they have forgotten. During the execution of actions and preparations of the next meeting, loggings may be made of information the participants consult and of questions they ask other meeting participants (for instance via e-mail) about the content of the meeting. At the end of the meeting cycle, the participants may be interviewed to ask their specific opinions on how they would like to have received support.

The user requirements that are extracted following the above non-obtrusive methods form the basis of developing the first concepts of a meeting browser. This (initially low-tech) concept can be introduced in the next meeting cycle to evaluate its usability and to gather recommendations for improvement of the design.

4.2 Experimental setting

In an experimental setting, comparisons can be made of several aspects. By providing similar teams with exactly the same task, we can compare, for example:

- Different means: Meeting means have been frequently an object of study, such as in Short et al. (1976) and, extensively, in McGrath and Holingshead (1994) in which different types of conference systems are discussed, in Detienne et al. (2004) in which face-to-face and technology mediated design project teams are empirically compared, and in Sundholm et al. (2004) that investigates how the lay-out of meeting rooms supports co-located collaborative work.

- Different methods: Methods refer to prescriptions of how to do a particular task. There are studies that deal with the meeting method in general (e.g., Piët (1990)), and design method in particular Pahl and Beitz (1996);Kroonenberg, van den and Siers (1993), including theoretical comparisons between design methods (Roozenburg and Eekels (1998)).

- The impact of individual and team factors

- Different organizational structures: Reorganizations occur frequently. The paradigm could be used to compare old situation with a proposed new organizational structure.

- Environment: The optimal structure of the organization depends heavily on the predictability of the environment. Hierarchical organizations are well-suited for predictable environments, while dynamic markets demands self-managing teams. What might be the best structure for slowly changing markets?

A particular experiment we are currently planning is on determining the optimal size and configuration of design teams. A long-lasting discussion in the literature is about the effectiveness of brainstorming. Punte and Hamel (2001) states that the exploration of the solution space of an individual designer is restricted, but can be enlarged when the designer cooperates with co-designers. In contrast, the amount of acts within a team grows exponentially with the number of participants, subsequently decreasing the efficiency of the process (Shaw (1971)). We will experimentally compare teams in which the participants have the roles as described in the example above with teams in which the industrial designer and user interface designer operate as co-designers.

5 Discussion and Conclusion

The paradigm may be used to investigate other types of tasks than design: i.e., crisis management, software development, and planning & control meetings. While most of the infrastructure remains the same, other input factors may be relevant. Of course, a new scenario needs to be developed, which means that new roles, events, and information bases have to be specified.

One question that remains is how real the behavior is that will be generated with the meeting paradigm. This is will only be known after the first try-outs. However, according to Driskell and Salas (1992), studying design teams in contrived settings remains valuable for theory development. They discuss the critique on artificial team settings that are so far away from real team behavior, and come to the conclusion that the advances in understanding team performance are most likely to result from the dynamic interplay between theory, research that test theory, and applications.

In conclusion, we have described a meeting paradigm, and are in a process of applying this paradigm in exploratory and experimental settings. Results from these studies may lead to adjustments of the paradigm. Ultimately, the paradigm will result into a theory of meeting behavior, which can help to improve the efficiency of organizations.

References

Chiu, P., Boreczky, Girgensohn, A. & Kimber, D. (2001). LiteMinutes: an internet-based system for multimedia meeting minutes .In: *Proceedings of WWW10, May 1-5, Hong Kong.*

Detienne, F., Boujut, J. & Hohman, B. (2004). Characterization of Collaborative Design and Interaction management Activities in a Distant Engineering Design Situation. In M. Zacklad et al., editors, *Proceedings of Cooperative Systems Design,* pages 83-98, 2004.

Driskell, J.E. & Salas, E.. (1992). Can you study real teams in contrived settings? Group research to understand teams. In E. Salas & W. Swezey (Eds.), *Teams: their training and performance.* (pp. 101-123). Norwood, NJ: Ablex.

Guillemot, M. Wellner, P., Gatica-Pérez, D. & Odobez, J-M. (2003). A hierarchical keyframe user interface for browsing video over the internet. In: *Proceedings of the 9^{th} IFIP International Conference on Human-Computer Interaction INTERACT 2003, Zürich.*

Guzzo, R. A. & Salas, E. (1995). *Team Effectiveness and Decision Making in Organizations.* San Francisco: Jossey-Bass Publishers.

Kroonenberg, H. H. van den & Siers, F. J. (1993). *Methodisch ontwerpen.* Culemborg: Educatieboek.

Lisowska, A. (2003). Multimodal interface design for the multimodal meeting domain: preliminary indications from a query analysis study. IM2.MDM Report 11, ISSCO/TIM/ETI, Université de Geneve.

McGrath J. E. & Holingshead, A. (1994).*Interacting With technology: Ideas, Evidence, Issues and an Agenda.* Thousand Oaks: SAGE Publications.

Pahl, G. & Beitz, W. (1996). *Engineering design; a systematic approach.* London: Springer.

Piët, S. (1990). Overleg, vergaderen en onderhandelen. Groningen: Wolters-Noordhoff.

Punte, P.A.J., Hamel, R.. (2001) *Cognitieve beschrijving van het ergonomisch ontwerpen [Cognitive description of the ergonomic design process].* TM-rapport nr: TM-01-D005. Soesterberg: TNO Technische Menskunde

Roozenburg, N.F.M. & Eekels, J. (1998). *Productontwerpen, Structuur en Methoden.* Utrecht: Uitgeverij Lemma BV.

Shaw, M. E. (1971). *Group Dynamics: The psychology of Small Group Behavior.* New York: McGraw-Hill Book Company.

Short, J., Williams E. & Christie B. (1976). *The Social Psychology of Telecommunications.* New York: John Wiley and Sons.

Steiner, I. D. (1972). *Group process and productivity.* New York: Academic press.

Stewart, G. L., Manz, C. C. & Sims, H. P. (1999). *Teamwork and group dynamics.* New York: John Wiley & Sons, Inc.

Sundholm, H., Artman, H. & Ramberg, R. (2004). Backdoor Creativity – Collaborative Creativity in Technology Supported Teams. In M. Zacklad et al., editors, *Proceedings of Cooperative Systems Design,* pages 99-114, 2004.

Swigger, K., Brazile, R., Peng, X. & Harrington, B. (2004). Computer-Supported Collaboration and the Effects of Culture. In Darses et al., editors, *Supplement to the Proceedings of Cooperative Systems Design,* pages 13-20, 2004.

Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the Unexpected: Assuring High Performance in an Age of Complexity.* San Francisco: Jossey-Bass.

Getting used with groupware – a *First Class* experience

Dulce T. Pumareja, Klaas Sikkel
Information Systems Group
Department of Computer Science
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
P.O. Box 217 7500 AE Enschede, The Netherlands
{pumareja|sikkel}@cs.utwente.nl

Abstract

This paper reports an empirical investigation of a long-term use of a groupware system in a spatially and massively distributed network of educators. It is a case study based investigation aimed at understanding the impacts of collaboration technology in supporting social interaction. The paradigm of social constructivism and the perspective of *structuration* are proposed as frameworks for understanding the impacts of technology on mediating social interaction. Utilizing these perspectives in an empirical investigation, the case study findings demonstrate how collaboration technology can serve as a change agent in transforming the culture and structure of social interaction. This is enacted in two ways: through the various meanings which people construct when interacting with technology and in benefiting from the structural properties of a system through its technical *affordances*.

Keywords: Groupware, Structuration, Implementation, Social Constructivism, ICT in Education, Case Study

1 Introduction

Groupware is a general term for a repertoire of ICT applications that support cooperative work between and among groups of people. Collaboration technologies, cooperative systems, coordination tools, group support systems, etc. are synonymous with groupware technology. Groupware systems are used to support and carry-out the social domains of work, i.e. the aspect of one's job in which interpersonal interaction and cooperative processes take place. These are activities that are basic to most tasks but are not officially specified as part of one's tasks. For example, one sends an email to a colleague asking a favor if she can get a copy of the presentation he gave the other day because it contains a nice presentation template that she can use for her report. Actions like these are informal, intermittent, synergetic and largely unstructured. These properties make them difficult to specify in advance. Consequently, these are the activities groupware systems are meant to support. According to Andriessen (2003), groupware systems can be distinguished from other ICT applications by having functions that serve the following human interaction processes:

- communication, i.e., exchanging signals
- cooperation, i.e., working together, making decisions
- coordination, i.e., adjusting the work of group members, leadership
- information sharing and learning, i.e., exchanging, sharing information and knowledge
- social interaction, i.e., group maintenance activities, developing trust, cohesion, conflict handling, reflection.

By mediating human interaction and communication processes, groupware systems have the potential to bring about dramatic changes to the social functioning of individuals, groups and organizations. Extensive research has been conducted on the impact of groupware on group performance, dynamics, outcomes, as well as organizational functioning (Galegher and Kraut 1994; Orli 1992b; Vogel and Nunamaker 1990). For example, Bikson & Eveland (1990) and Connolly (1996) report that the anonymity feature of a group decision support system enlarges the scope of user participation and enables the expression of negative opinions. In normal face to face meetings such are rather difficult to convey. In return, anonymity helped improve the quality of input to group decision-making. In another study, an integrated groupware system has been observed to bring about changes to the way people collaborate in the form of increased pro-activity, increased utilization and dissemination of knowledge and new forms of coordination (Orli 1992a). While there have been numerous studies that report the positive consequences of groupware technology, a number of studies have also observed instances of groupware implementation failures (Pumareja et al. 2003; Grudin 1988). These failures relate to issues such as user acceptance, adoption, organizational and cultural contexts and poor system design (Andriessen 2003; Grudin 1995).

In the following case study, we try to expand the scope of empirical research aimed at understanding the impact of collaboration technology on social interaction. Utilizing a social constructivist paradigm and a *structuration* perspective as theoretical take-off points for analysis, the investigation of the case is guided by the following questions:

- How does a spatially distributed group of workers make sense of groupware technology?
- How do the collaborative features of a groupware system shape distributed social interaction?

The case study was conducted among a network of geographically dispersed educators working for an institute of higher education. Qualitative research approaches were employed in conducting the case investigation – in-depth interviews, demonstrations of system use, system inspection and participation in using the system, and document analysis. A total of 17 interviews were conducted from a representative group of educators (14) and system administrators (3). These respondents were visited in their place of work, which is mostly at home and in the central and regional offices. They were asked to demonstrate how they use the system and their digital workspaces were seen and observed.

The presentation of the findings of this study is structured as follows. We first discuss the theoretical framework followed by the research setting and work context, the presentation of the case and its analysis and a concluding discussion.

2 Theoretical framework

Social constructivism is one of the dominant paradigms used to analyze and understand impacts of groupware technology into the social domain of users (Bijker 1990; Fulk 1993; Mackenzie and Wajcman 1985). This thinking embodies the notion of a social shaping of technology through the shared construction of meaning, interpretation and sense-making by human actors. Within this perspective, technology is considered to be equivocal (Wieck 1990). It can achieve multiple meanings which could be both consistent and conflicting, despite the fact that there is only one possible technical description through its specification. Drawing from this paradigm are several theories of technology and social interaction that aim to provide more profound conceptual lenses for studying the impact of groupware technology. One example of these is the *structuration* perspective, which takes root from Giddens' (1984) meta-theory of *structuration* (Orli 1992b; Poole and DeSanctis; Hettinga 2002; Ruël 2001). The main ideas of *structuration* theory as synthesized by several authors (Hettinga 2002; Clarck et al. 1990; Volkoff 1999; Andriessen 2003) are:

- Structures consist of both social entities and their properties, i.e. organizations, groups are social structures as well as their hierarchies, rules and resources; a groupware system in an organization with its hardware and software components is a structure in the same way

an institutionalized system such as taxation with its own rules and procedures is also an instance of a structure.

- Structures serve both as medium and outcome of action. Humans draw upon structures for action which could enable or constrain, and in return reproduce them (structures) as a by-product of interaction. *Structuration* is the process of producing and reproducing (reinforcing) structures in a dynamic interaction between humans and social systems. This is called the principle of duality of structures.

- Human agents are knowledgeable and are capable of exercising their powers to accomplish a social practice, i.e. they can choose to change or confirm structures.

- There are three modalities of *structuration*, i.e. domains in which the interaction between human action and social structures take place. These are in the domains of meaning-constitution, power relations, and norms and legitimation.

Applied in the context of technology, Orlikowski (1992b) extended this perspective into a *structuration* model of technology. She put forward the notion of duality of technology – technology is used as a means to accomplish some action and is likewise created and changed by human action. This is reflected in the two modes of human interaction with technology, namely, design and use. Technology is physically constructed by actors in a given social context and is socially constructed by actors through the different meanings they attach to it and in the various features they emphasize and use (Orli 1992b). When making use of a groupware technology, users draw upon its features and resources (structures) such as enabling communication across time and space. The continued habitual use of this medium reinforces the structural role of the technology in the organization that implemented it. Over time, the groupware system becomes institutionalized as the medium of communication of the organization, making office memos and bulletin boards obsolete. The institutionalization of technology is its detachment and objectification from human action, thereby becoming a structural property of the organization. The *structuration* model of technology is also premised on the knowledgeable and reflexive aspects of human actions. That is, agency is the capability to perform action which can have intentional and unintentional consequences. In a related study, Orlikowski (1996) studied the use of Lotus Notes where it was implemented in consulting organization with the intention of supporting collaboration and knowledge sharing. However, the competitive culture of the organization had resulted into a non-collaborative use of the system. It was instead used a personal productivity tool.

The *structuration* perspective of technology fits well with the notions of improvisations, appropriation, innovations and socio-technical evolution that are associated with groupware implementations. Groupware implementations have been observed to undergo a 'drift', where the original intentions for use differ from the actual exploitation and purposes in which it is used for Ciborra (1995).

3 Research Setting and Work Context

The case study was conducted among a distributed group of educators who provide training and consulting in the domain of special education and orthopedagogy. This network of educators makes up the teaching faculty of the Institute for Orthopedagogy. The Institute is one of the departments of the Faculty of Education of large city college in the Netherlands.

The Institute for Orthopedagogy is one of three institutes of higher education that provide postgraduate level education in the country. Therefore it operates at the national level, with the head office in Utrecht and regional offices nationwide. Its core line of business is the provision of postgraduate education and training for teachers who would like to earn a qualification diploma in order to teach special education. Teaching special education is a regulated profession, meaning a successful completion of bachelors-level education in teaching is necessary prior to enrolling in the course. Next to this, it also provides in-house training and consulting services to primary schools nationwide that have a special education component. Its clients include most primary education teachers, special education teachers, primary schools and to a certain extent private

individuals. Special education is generally understood as educating students, in this case primary school children, with special needs, i.e. autistic, dyslexic, children with handicaps – blind, deaf, mute, children with learning disabilities among others. From time to time, the school also performs clinico-psychological diagnosis of children with mental and emotional problems. The background of the educators is diverse – some of them are psychologists, teachers with master specializations or teachers with extensive teaching experience.

The provision of postgraduate education and training is organized in such a manner that the teachers (the students of the school) receive instruction at the place where they live and work. The spatial distribution of the educators across the country is a consequence of this organizational set-up and work configuration. Most of the educators work at home. The Institute provides them with a home-office facility that includes furniture, computer and software. The educators would then have to provide a space in their homes as work area. The work is structured in a flexible manner, not an 8-5 schedule. Sometimes, an educator finds herself giving a 3 hour lesson in the morning or in the evening, or attending a meeting with the administrators of the school. When not attending meetings, giving lessons or having appointments, most of them perform their other tasks at home, i.e. curriculum development, lesson planning, checking papers, etc. When working at home, they have the freedom to structure their own schedules, i.e. having an appointment with the plumber during the day and compensating that by working late at night.

The execution of one's tasks is largely performed on an individual basis. Collaborative activities take place in between, i.e. curriculum and instructional development, course evaluation, planning and coordination, etc. Most of these activities take place through the groupware system. The central office together with the competency centers which each educator belongs, also organize several face to face meetings to further stimulate collaborative activities. However, in most instances, the educators work alone. In some locations, an educator hardly sees her colleagues in the region nor is she familiar with them. But that does not mean that she gets the feeling that she does not know them or cannot be acquainted with them. The groupware system helps her overcome that obstacle.

4 Background to the Acquisition and Implementation

Eight years ago, the ICT Manager and a core group of educators were acquainted and impressed with the experiences of similar institution from Sweden in implementing a groupware system. The Swedish institute was using a Macintosh-based groupware for supporting its teaching and learning processes. The Macintosh compatibility of the system augured well for the Institute, who at the time was using those kinds of machines. Further, the Institute was keen on applying ICT applications for educational and support purposes in the light of the increased diffusion of information technology. Providing the motivation for this is the generally shared view in the educational sector that regards the use of ICT applications positively. This is inline in bringing about renewal and improvements in the quality of education for which massive support and funding was available.

The collaboration system that the Institute acquired and implemented is an integrated COTS (Commercial Off The Shelf) groupware application called First Class Client (Open Text Corporation). The users usually referred to it as FCC for brevity. It was originally built for Macintosh machines. Its collaborative features were mainly marketed in the beginning as support environment for online learning. Among its very first users were schools and educational institutions. Its design concept and intention as tool support for online learning continue to hold despite having been proprietarily bought by a third-party software development company. It remains to be associated in the domain of educational and learning support.

As an integrated collaboration system, First Class Client contains a range of built-in features for enabling collaboration. These are email, shared workspaces, asynchronous chat, conferencing, individual and shared calendars, private and public directories. Further, the system has an accompanying programming environment for developing customized applications.

The system was bought in 1995 and was put into pilot use among 23 educators living within the environs of the central office. Later it was implemented with the secretaries in the main office. This

process was administered by a former educator who later became the administrator for the system together with other technology-savvy educators keen about the system. These technology keen educators saw the bigger potential of the system as a communication infrastructure for overcoming the challenges of a distributed organizational set-up. These people became project champions for the eventual roll-out of the system throughout the Institute, effectively replacing the existing email system. The implementation was supported by mandated policies for use which were integrated with the home-office package support provided by the Institute for its employees. Most of the computer units acquired for the home office were IBM compatible PCs, hence the Institute also switched to the PC version of the system which was already available.

5 Features of Long-term Groupware Use

The deployment of First Class Client enabled a series of changes in the Institute particularly in terms of facilitating communication and social interaction as well as transforming work. The long-term implementation of the system can be classified into two phases of use: early use phase and an established use phase.

5.1 Early use phase

The early use of the system in the Institute is a period of adjustment for getting used with the system. It is characterized by efforts aimed at promoting the use of the system.

Positive and opportunistic attitudes towards technology

The implementation of the system was largely brought about by a positive and opportunistic attitude towards technology of the small group of people responsible for the IT resources of the Institute. This was reflected in their tech-savvy mindsets that eventually made them the system's sponsors for the entire Institute. These were the people who have made their acquaintance eight years ago with the Swedish educational institute that was using First Class Client. Their positive mental model of technology was reflected in their profound appreciation of the collaborative features of groupware technology which they thought is something the Institute needs.

Institutionalization

The profound appreciation of the possibilities of groupware technology was influential in establishing an institutionalized use of the system. This was enacted through an official mandatory use policy of the system. The system administrator, being in a powerful position as IT resources manager and having advanced knowledge about FCC, played a key persuasive role in formulating and executing the policy of compulsory use. All educators working for the Institute had to make use of the system. They were given installation CDs for their home PCs. Every new employee who was entitled to a home PC would get the unit with a pre-installed system. At the same time, they have two weeks to get familiar with the system and be online.

Drift in intended use

While the core design concept of FCC was to promote an online learning environment, i.e. support the communication between teachers and students, the system sponsors in the Institute have instead chosen to use the system to support internal communication in the Institute. The system was utilized as a communication infrastructure for bringing together the geographically distributed educators. This action is a by-product by their opportunistic outlook about the possibilities the system and relating it to the current problems encountered by the educators. This is reflected by the system administrator:

"Before, the teachers had to drive all the way to Utrecht to bring their materials for reproduction. Or they send it to each other by post and fax. A portion of their telecommunication expenses is reimbursable. With FCC, all of that has become easier and faster."

Technological accommodation

Despite getting no official training on how to use the system, the educators however compensated for the system. That is, they spent extra time to learn about the system as well as taking efforts in asking for help. This is attributable to the fact that users do not have a choice – it is the only computer-based communication tool available next to the telephone and they are mandated to use it.

5.2 Established Use Phase

As system use intensified, with more users becoming part of the online network, a series of changes were observed, and the system took a more defined role in the Institute.

New forms of communication and collaboration

Communication and collaboration in the Institute took a massive transformation with the implementation of FCC. All educators agree that with FCC, communication and collaboration have become more efficient, i,e. faster and more convenient. Whereas before, most of them would send each other material by post and fax, they now send each other digital files through FCC. Time savings were realized through FCC with the capacity for accomplishing more tasks enabled. Likewise, it enabled them to communicate and collaborate with colleagues from other locations whom they do not even know. It is also now possible for them to send messages to a large group of people at the same time. The system also facilitated the knowledge-sharing among educators in ways not previously possible. One educator remarks:

"When I have a question or I need to some references about a topic, say dyslexia, I just send an email in that particular knowledge center or to my group and I get replies right away."

'Closed system' mental model

The convenience of not having to remember a colleague's email through the public directory feature of FCC had led to a closed system mental model of FCC. The users have formed a rather consistent and shared view of FCC as an exclusive system. They thought that for other people outside the Institute, it is not possible to send email to or receive email from them. However, in fact, the system allows for this technical functionality. All they need to simply type the email address of the person they would like to send email to. Only because they cannot find the names of these people in the public shared directory, they thought it is not possible. As a consequence, most educators maintain a second email address from free email services such as Hotmail and other internet service providers. The following educators' comments give insight to this:

"*When I went to Malaysia for vacation last winter, I was able to access FCC through the web. I can read my email and get in touch with my work. However, I applied for a Hotmail account so that I can send email to my children. But I didn't know that it is possible to do that via FCC. However, my children are not part of FCC, they are not in the system and they don't have it."*

"*I also have another email address. I use that to communicate with the clients. I do that because they are not on FCC and they do not have FCC*."

Likewise, this mental model of a closed system is also shared by the system administrator. Before this research was undertaken, the system administrator suggested that it is better for the researcher to be part of FCC in order to gain access to the educators. An account was created for the researcher for the purpose of accessing and coordinating with the users.

Socialization Platform

In this later period of use, the system had become the de-facto socialization space, especially for new employees. This is the space where new employees get to know their colleagues and introduce themselves. This is because the set-up of the work in the Institute does not allow for the usual new employee introduction and socialization. This is reflected in the comment of a newly-hired educator of the Institute.

"I met with the regional manager and the first thing he asked me was, 'are you already on FCC? You know, the coordinator of the knowledge center you belong also lives here in the area. You can find her name in the address book and maybe it is useful to introduce yourself to her'."

Norms for interaction, social responsibility and forced reciprocity

In the course of continued use of the system, informal norms for interaction emerge. These norms were formulated out of an unconscious process, i.e. it was not planned. Implicitly, everybody is expected to respond promptly to messages as a form of courtesy. Users were pleased and at the same time surprised at the prompt replies of their colleagues when they ask a question or request information. In return, they get the feeling that they are obliged to reply back as soon as they can. This way, the system evokes in them a sense of forced reciprocity.

Sense of urgency

The features for supporting communication not only enabled faster communication among the educators but it also induced a continued sense of urgency. With colleagues responding quickly and promptly, some educators felt the increased pressure of work.

"I find the system rather stressful. It increases work pressure unnecessarily. It gives you the feeling that work does not stop."

Balancing individual privacy and work responsibility

On the other hand, for some educators, they find the system supportive in striking a balance between their individual private life and work responsibility. The system afforded them the convenience of receiving messages when they are not at home or responding to a message at the time when it is convenient.

"*You see, with the telephone, you have no choice. You have to pick it up right away when it rings. But with FCC, you can use it when it is convenient for you."*

Changed work rituals and habits

The system has also brought about changes in work rituals and habits among the educators. For example, it had become a ritual for most of them to go and check their email first thing in the morning or before going to a meeting for any last minute changes. Likewise, work patterns have changed, i.e. working in the weekends and late at night. Some of them even cannot go to bed without checking their email for the last time. While these changes in work patterns are largely attributable to the nature of work in the Institute, the system however, forges a reinforcement of such.

System as the conceptual representation of the organization

In the absence of a physical space symbolizing the organization within their sight, the users have turned to the system and their computers for orientation. In other words, the system serves as their reference for anything related to their work because they are distributed in space. It is their window to the organization they work for. One educator puts it:

"That little FCC icon on the desktop for me represents the Institute."

Reflection and compassion

Characterizing the later implementation and its sustained use is the continued reflective activities shared by the educators relating to the system. When they have the occasion for face-to-face meetings, they talk about the problems they encounter with the system. Altogether, they make suggestions on how they can improve their skills in using the system better and more efficiently. For example, a group had asked the system administrator to provide training for them. They also ask each other how they perform certain tasks with the system when the system administrator is not available.

Reflection concerning the system at the individual level is also observed. This led to an increased appreciation of the value of the system and to enthusiasm in sharing this with others outside the Institute. As one educator puts it:

"You see I also consult for a series of primary schools. I heard that there are plans of setting up a digital knowledge network for the teachers. I would like to advise them to consider using FCC."

Likewise, they are also concerned when one is having problems with using the system, i.e. someone is not used to using computers. One particular educator manifests her compassion with her colleagues this way:

"*When I make a reply or send a forwarded message, I remove the unnecessary information on the text that is not relevant. I find that pollution and not necessary. We already had a lot of work to do and we do not have time to sort all information."*

Institutional inevitability and invisibility

Over time, the FCC had become an inevitable tool for the Institute. The cooperative aspect of the work of the educators is effectively supported by it, such that had become invisible in the communication and collaboration process:

"I use FCC nowadays more than I use the telephone."

"*When I want to ask a colleague something, my first instinct is to use FCC. Unless it is something that needs to be discussed more thoroughly, then I call or we make an appointment for a meeting."*

6 Discussion

The findings of the case analysis demonstrate how technology can serve as an agent of change in transforming the culture and structure of social interaction in a distributed organization. The notion of a mutual shaping of technology and social structure encapsulated by the *structuration* framework is useful in gaining an insight of the mediating effects of collaboration technology to social interaction. These notions are operationalized in terms of two analytical frameworks:

- Shared construction and assignment of meaning to collaboration technology
- Shaping of social interaction through the technical features and *affordances* of collaboration technology

6.1 Shared construction and assignment of meaning to collaboration technology

As early as the planning and initial use of the system in this case, social construction and assignment of meaning is already in place. This was manifested in the taking advantage of the collaborative functionality as a communication infrastructure. The system sponsors did not proceed with the intended use of the system as an online learning environment. Instead, they implemented the system as a communication infrastructure for bringing together the spatially dispersed educators. This decision is brought about by their reflective examination of the problems they encounter as a distributed organization where communication and effective performance of tasks were time-consuming, inefficient and sometimes irritating. Through their widened mental model of technology indicated by their profound appreciation of the collaborative features of groupware technology, the systems sponsors had constructed the system as a solution to their problem.

In the later established use of the system, various mental models of the system were evoked among the users as a result of interacting with it. First among these is the mental model of a closed and exclusive system where the system, unintentionally, had become instrumental in drawing the boundaries of social interaction for the group. For the network of educators, it was clear to them who are part and not part of their social network on the basis of who are the people they can see as listed in the public directory. At the same time, people reify their membership and make their presence visible in the social network by making use of the system to participate in the

online interaction. This creates another level of social bounding which distinguishes those people who are active online and those who are not. Next to this is the mental model of a window to the organization. Despite that work is executed largely on an individual basis, the supporting processes and information needed in effectively accomplishing the tasks have a social context. By being distributed in space, the educators felt isolated in their work and they turn to the system for orientation and for keeping in touch with the organization and their colleagues.

6.2 Shaping of social interaction through the technical features and affordances of collaboration technology

Independent of the meaning users assign to a system, the collaborative features of a technology in interaction with its users also shape the structure of social interaction. Likewise, it can also bring about unintended consequences to the social environment. These are known as the *affordances* of technology.

Becoming a virtual space for socialization is a foremost example of the technological *affordances* of the groupware system. It supplanted the physical space where socialization takes place into something symbolic and virtual.

The implementation of FCC in the Institute also led to the formation of several structures and behavioral changes. These are in the form of norms, i.e. making agreements, replying promptly, new culture of collaboration and communication, such as increased social responsibility. On an individual basis, the system brought about changes to the work habits of people by making the system part of their work ritual. The assimilation of technology into their work in return affects the other people in the network. Giving prompt and quick reply, or replying late a night to the email of a colleague gives the recipient the feeling of forced reciprocity, a sense of urgency when not necessary, and stress.

Consistent with the *structuration* perspective, the presence of reflexivity is also observed in this case. Reflexivity is found in the shared effort of the users to talk about the system. FCC was not only the medium of communication for the users, but it is also the object and subject of communication for them. In other words, the system was used to mediate communication about itself. Consequently, this had helped in sustaining a positive view about he system, appreciating its usefulness and putting it into continued use.

Lastly, it is also observed that when a collaborative system has succeeded in becoming a useful mediation tool, the tool becomes invisible. It is assimilated into the communication and collaborative processes of the users, such that when it breaks down, the social interaction also breaks down.

References

Andriessen, J.H.E. (2003). Working with groupware: understanding and evaluating collaboration technology. *Springer Verlag*, London, 2003.

Bijker, W. (1990). The social construction of technology. PhD Thesis. University of Twente, *Alfa*: Enschede.

Bikson, T., and Eveland, J.D. (1990). The interplay of work group structures and computer support. In R. Kraut, J. Galegher, C. Egido (Eds) Intellectual Teamwork. *Erlbaum Associates*, Hillsdale, NJ, pages 245-290.

Ciborra, C.U. (1995). Introduction: what does groupware ean for the organizations hosting it? In Ciborra, C.U., Groupware and teamwork: invisible aid or technical hindrance? *John Wiley & Sons*, Chichester.

Clark, J., Modgil, C., and Modgil S. (1990). Anthony Giddens: consensus and controversy, *Falmer Press*.

Connolly (1996). Electronic brainstorming: science meets technology in group meeting room. In S. Kiesler (Ed.) Research Milestones on the Information Superhighway, *Social Science Research Council Press*, New York.

Fulk, J. (1993). Social construction of communication technology. *Academy of Management Journal*, 36(5), pages 921-950.

Galegher, and Kraut (1994). Computer-mediated communication for intellectual teamwork: an experiment in group writing. Information Systems Research, 5(2), pages 110-138.

Giddens, A. (1984). The constitution of society: outline of the theory of structuration. *University of California Press,* Berkeley, CA.

Grudin, J. (1988). Why CSCW applications fail: problems in design and dvaluation of organizational interfaces. In *Proceedings of the Conference on Computer Supported Cooperative Work 1988 (CSCW '88)*, ACM/SIGCHI & SIGOIS, NY, pages 85-93.

Grudin, J, and Palen L. (1995). Why groupware succeeds: discretion or mandate. In Proceedings of the Fourth *European Conference on Computer Supported Cooperative Work* (ECSCW '95), Kluwer, Dordrecht, The Netherlands, pages 263-278.

Hettinga, M. (2002). Understanding evolutionary use of groupware. PhD Thesis. *Telematica Instituut*, Enschede The Netherlands.

Mackenzie, D. and Wajcman, J. (1985). The social shaping of technology. Open University Press, Milton Keynes, UK.

Open Text Corporation. http://www.firstclass.com.

Orlikowski, W.J. (1996). Evolving with Notes: organizational change around groupware technology, in Ciborra, C. (Ed) Groupware & Teamwork*, J. Wiley*, Chichester et al. 1996, pages 23 - 60.

Orlikowski, W.J. (1992a). Learning from Notes: organizational issues in groupware implementation. *In Proceedings of the Conference of Computer Supported Cooperative Work (CSCW '92),* ACM Press, pages 362-369.

Orlikowski, W.J. (1992b). The duality of technology: rethinking the concept of technology in organizations. *Organization Science*, 3(3), pages 398-427.

Poole, M.S. and DeSanctis, G. Understanding the use of group decision support Systems: the theory of adaptive structuration. In J. Fulk & C. Steinfeld (Eds*),* Organizations and Communication Technology, *Sage Publications*, Newbury Park, pages 173-193.

Pumareja, D.T., Bondarouk, T., and Sikkel, N. (2003). Supporting knowledge exchange isn't easy: lessons learned from a case study. *14^{th} Information Resources Management Association International Conference*, Philadelphia, May 18-21, 2003.

Ruël, H. (2001). The non-technical side of office technologies! PhD Thesis. University of Twente, *Twente University Press*, Enschede, The Netherlands.

Vogel, D. and Nunamaker, J. (1990). Design and assessment of a group decision support system. In R. Kraut, J. Galegher, C. Egido (Eds) Intellectual Teamwork. *Lawrence Erlbaum Associates*, Hillsdale, NJ, pages 511-528.

Volkoff, O. (1999). Using the structurational model of technology to analyze an ERP implementation. *Americas Conference on Information Systems 1999*, August 13-15, Milwaukee, WI.

Weick, K. (1990). Technology as equivoque: sense-making in new technologies. In P.S. Goodman, L.S. Sproull & Associates (Eds.), Technology and Organizations, Jossey-Bass, San Francisco, pages 1-44.

Psychologically Intelligent Mobile Multimedia Messaging Systems

Timo Saari ♣ ♢ saari@hkkk.fi
Marko Turpeinen ♣ marko.turpeinen@hiit.fi
Jari Laarni ♢ laarni@hkkk.fi
Niklas Ravaja ♢ ravaja@hkkk.fi
Kari Kallinen ♢ kallinen@hkkk.fi

♣ Helsinki Institute for Information Technology, Finland
♢ Center for Knowledge and Innovation Research, Finland

1 Introduction

Mobile messaging is an increasingly important way of social interaction as people use their mobile phones for communicating with each other with textual and multimedia messages. However, what is naturally often lacking in the communication interactions is the fine-tuning of the communication and the cues present in face-to-face communication. This may result in breaks and misunderstandings in communication. Another issue is the need of users to communicate their own emotions or facilitate a given emotion in the receiver of their message with limited means. This paper will explore the use of an information personalization system that may enhance cognitive processing and emotional communication especially in mobile multimedia messaging systems from the point of view of psychological theory thereby making such systems "psychologically intelligent".

Media- and communication technologies, such as systems for mobile messaging, as special cases of information technology may be considered as consisting of three layers (Benkler, 2000). At the bottom lies a *physical* layer that includes the physical technological device and the connection channel that is used to transmit communication signals. In the middle is a *code* layer that consists of the protocols and software that make the physical layer run. At the top is a *content* layer that consists of multimodal information. The content layer includes both the substance and the form of multimedia content (Billmann, 1998; Saari, 2001). Substance refers to the core message of the information. Form implies aesthetic and expressive ways of organizing the substance, such as using different modalities and structures of information (Saari, 2001). Naturally, these are difficult to separate.

Technologies are most often designed from the point of view of available communication capacity, software and hardware around a certain task. In addition to this there is another approach to design of technology taking into account the user experience of the users of technology and the goals users may have regarding a certain technology. For instance, a user may wish to perform a certain task as efficiently as possible, or a user may wish to be in a good mood when performing the task. Hence, there is another layer for design of technology in addition to the task-based approach.

This type of design may be called as being Mind-Based, i.e. paying attention to the needs and goals of users to achieve desired user experiences, such as positive mood and efficiency of cognition. It also involves personalization and customization that entail the automatic or semi-automatic adaptation of information per user in an intelligent way with information technology (see Riecken, 2000; Turpeinen, 2000).

Media- and communication technologies may be called Mind-Based if they simultaneously take into account the interaction of three different key components: i) the individual differences of

perceptual processing and sense making of different segments of users, ii) the elements and factors inherent in information and technology that may produce psychological effects (physical, code and content layers), and iii) the consequent transient psychological effects emerging based on perception and processing of information at the level of each individual. (see Saari and Turpeinen, 2004). Hence, Mind-Based design takes into account individual differences in processing information in order to be able to offer a particular user a particular type of experience. Naturally task and for instance the content or substance of information, such as email, Mobile Multimedia Messaging (MMS) or a news article, influence the user experience.

However, it may be feasible to facilitate desired user experiences just by varying the form of information. For instance, with Mind-Based Technologies one may vary the form of information per user profile, which may systematically produce, amplify, or shade different psychological effects (Saari, 2001; Saari, 2002; Saari, 2003a, Saari, 2003b). This type of system design approach may be of practical use, as it is known that individual differences in processing information may produce sometimes quite large variance in the intensity or type of psychological effects, such as depth of learning, positive emotion, persuasion, presence, social presence and other types of psychological states and effects (Saari, 2001; Saari, 2002; Saari, 2003a; Saari, 2003b).

Hence, the Mind-Based Technologies- approach may be valuable also when creating systems for facilitating clear and cognitively optimised or emotionally focused social interaction system. One operationalization of Mind-Based Technologies in system design is Psychological Customization (Saari and Turpeinen, 2004; Turpeinen and Saari, 2004).

2 Systems for Psychological Customization

Psychological Customization includes modeling of individuals, groups, and communities to create psychological profiles and other profiles based on which customization may be conducted. In addition, a database of design rules is needed to define the desired cognitive and emotional effects for different types of profiles. Once these components are in place, content management technologies can be extended to cover variations of form and substance of information based on psychological profiles and design rules to create the desired psychological effects.

The key idea is that matching a particular individual difference, such as user´s fluency in processing textual information, with a desired user experience, such as the need for efficient cognitive processing. This results the system presenting the user with as much text-based modality within a certain task as possible. This would then enhance information processing related to task of that particular user. Another user may be more fluent in processing audio information or video material with the same task to achieve more efficient processing of information, for example.

Psychological Customization can be applied to various areas of HCI, such as Augmentation Systems, Notification Systems, Affective Computing, Collaborative Filtering, Persuasive Technology and Messaging Systems. It can be hypothesized that the selection and manipulation of substance of information takes place through the technologies of the various application areas of Psychological Customization. Underlying the application areas is a basic technology layer for customizing design. This implies that within some limits one may automatically vary the form of information per a certain category of substance of information. The design space for Psychological Customization is formed in the interaction of a particular application area and the possibilities of the technical implementation of automated design variation (see Saari and Turpeinen, 2004)

The particular focus of this paper is messaging systems, specifically mobile multimedia messaging. Even though no actual system has been implemented yet for Psychological Customization related to mobile messaging content presented on small screens, empirical evidence supports the feasibility and validity of this idea. First, it is well established that there are individual differences in cognitive processes such as attention, memory and language abilities. These individual differences have a considerable effect on computer-based performance (e.g., Egan, 1988). For example, individual differences in memory capacity have an effect on people's behavior in many types of activities (Vecchi et al., 2001). This suggests the need for Psychological Customization Systems that optimize the presentation of information to different target groups having different psychological profiles. There is considerable evidence in literature that varying the form of information

creates for instance emotional and cognitive effects (e.g., Laarni, 2002; Laarni, 2003; Laarni et al., 2002).

In media studies it has been found that different modalities, such as visual and auditory, may lead to different kinds of psychological influences and the valence of a preceding subliminal stimulus influences the subsequent evaluation of a person evaluated (Cuperfrain and Clarke, 1985; Krosnick et al, 1992). In educational studies it has been shown that different ways of processing information influence learning and emotion of stimuli with certain modality (Riding and Rayner, 1998). Research concerning emotional influences on the cognitive processing of information has often concentrated on how different emotions related to information change the way users pay attention to, evaluate and remember the mediated message. This research has results on the influence of emotional information as increasing the user's self-reported emotion (Lang et al, 1996); attention (physiological and self-reported) (Lang et al, 1995) and memory for mediated messages, particularly arousing messages (Lang, 1990; Lang et al 1995; Lang et al, 1996). Studies in experimental psychology have shown that recognition and memory can be influenced or even enhanced by previous exposure to subliminal visual or auditory images of which the subjects are not consciously aware (Kihlström et al, 1992). Some of these effects are produced in interaction with individual differences, such as cognitive style, personality, age and gender.

In our own research on the influence of form factors of colour screen PDA´s and mobile phones of information (such as news, games, messaging content and entertainment content) on psychological effects has yielded many results. Regarding emotional responses has also shown that subliminal exposure to happy affective primes in connection with video messages presented on a small screen has several putatively positive influences (i.e., increased pleasure, perceived message trustworthiness, and memory; Ravaja, Kallinen, Saari, & Keltikangas-Järvinen, in press). Further, recent studies in our laboratory have empirically confirmed that media messages can be modified in terms of audio characteristics (Kallinen & Ravaja, in press; Ravaja & Kallinen, in press) and the presence of image motion (Ravaja, 2004) to meet the personality (as defined in terms of dispositional behavioral activation system sensitivity) of the user, thereby enhancing his or her attentional engagement, information processing, and enjoyment. The role of hardware should not be neglected. A device with a large screen or a portable device with smaller screen with user-changeable covers may also influence the emerging effects (e.g., Laarni & Kojo, 2001).

This empiric evidence partly validates the possibility for Psychological Customization Systems at least with mobile devices and user interface prototypes used in our own research. Typical experiments we have conducted on the influence of form of information on psychological effects have included such manipulations as animation and movement (for orientation response), fonts of text, layout of text, background colors of text, user interface navigation element shapes (round vs. sharp), user interface layout directions, adding background music to reading text, use of subliminal affective priming in the user interface (emotionally loaded faces) and use of different modalities of information, for instance.

3 Application for Mobile Messaging Systems

The motivation for mobile messaging is the need to communicate with another user of the system. The recipient of the message sent may be a friend or family member that makes the task of using mobile messaging partly related to staying in touch with one´s emotionally close individuals. The recipient may be a co-worker that asked for advice on a problematic situation. The task here may be to be able to communicate very clearly and avoid mistakes in the understanding the message.

It is obvious that in social interaction the users construct the substance of the messages. However, the form of the message may be varied with automated or semi-automated systems. The information needed to conduct these manipulations can be accessed via individual and social modeling and profiling of the users of the mobile messaging system. For instance, in a semi-automated system a user with an intention to create positive emotion in the other user with the way of presenting his textual message may utilize a background colour suggested to him by the system that predictably would induce a positive emotional state in the receiving user based on his profile. Similarly, cognitive efficiency may be created using a modality or layout that the recipient

most likely can process very fluently. A summary of the technological possibilities of customizing mobile messaging in a psychologically intelligent way is presented in Table 1. Traditional text-based SMS-messaging is presented as a comparison to the possibilities of multimedia messaging.

Layer of Technology	**"Traditional" Text-Based Mobile Messaging**	**Psychologically Intelligent Mobile Multimedia Messaging**
1. Physical	Small device, small black and white screen capable of displaying simple graphics and text. Button-based input.	Small device, larger colour screen. Various multimedia capabilities: text, audio, graphics, video etc. Camera, video recorder, audio recorder. Local networking possibilities: infrared, Bluetooth etc. Buttons, voice and pen as inputs.
2. Code	Menu-based user interface	Menu or windows- type of user interface.
3. Content **A. Substance** **B. Form** Modality	Fixed textual content written by the users. Fixed modality: text	Fixed multimedia content created by the users. Flexible modalities per substance, can be tailored to individual tastes. Text, audio, video, graphics, photographs, tactile aspects and animation may be used.
Visual presentation	Text, black and white over quite a small screen.	Flexible visual presentation of substance, i.e. different layout schemes are quite easy to make and download. Larger colour screen.
Structure	Linear text.	Linear text, video, audio, graphics, animation etc.
Way of interaction and user interface	Hardware and software controls to guide reading by scrolling and enable input.	Flexible ways of interaction based on technological possibilities: pen-based input, speech, hardware and software controls.

Table 1: Technological possibilities of psychologically intelligent mobile multimedia messaging systems.

In order to a Psychological Customization system to function in mobile messaging, the users would need to create a user profile (personality, cognitive style, other relevant information) for the system to gain access to form factors that may create desired cognitive or emotional effects in the users. The users would also fill out a community profile that indicates which users have authority to send psychologically intelligent messages to them and vice versa. The system would need a database of design rules of probable psychological effects of each type of manipulation per type of user and some other functionality. Further, if the system could record the user´s psychological state, this may make the system more reliable by making it possible for the system to more objectively verify the psychological states of the user.

Based on Table 1 it may be then said that Psychological Customization of Mobile Multimedia Messaging can be used in several ways with available technologies. First, in creating cognitive effects, the system may automatically manipulate the form of the message received or sent in a manner that is optimal for the receiver based on the rule-database. For instance, typography and screen layout may be optimised for each receiver. Also, when possible, automatic translations from text to audio or vice versa may be sensible. Second, for emotional effects, the system may offer the sender of messages graphical, video, audio or other types of MMS-templates to use in order to communicate a particular emotion of their own to the receiver or create a desired emotion in the receiver. In both cases the system would automatically suggest for the sender of message a possibility to psychologically customize the message for a particular receiver. The user would select a desired effect, such as creating positive emotion in the receiver with a message in which the substance is written in text and the system would present the sender with ready-made and psychologically evaluated templates (consisting of graphics, animation, sounds, videos etc.) that with high probability may create the desired emotion for the receiver with a particular user profile. The sender would type in the text-message in the template, finalize the design and then send the message. The receiver would receive an emotionally optimised message and may then experience the desired emotion. Naturally, if the substance of the message and the form of the message communicate a different emotion, for instance the substance is hostile and the template is joyful, some effects may not be realized.

If the capture of user´s psychological states with psychophysiological and behavioural methods is realized with future mobile technologies, it may be possible to more objectively capture the effects of a given MMS template on the receiver designed to induce positive emotion, for instance. Similarly, cognitive effects may be more traceable as well, perhaps related to the optimal amount of information to be presented, such as text on a small screen. Such recording technologies would make the system more reliable and allow for fine-tuning of effects and allow the system to learn each individual´s patterns of responses in order to conduct better Psychological Customization over time.

4 Conclusion

According to the authors´ knowledge no other comprehensive framework of varying form of information to systematically create emotional and cognitive effects has been presented, specifically in mobile messaging. Psychological Customization is based on the idea of providing targeted added value to the user, such as enhanced depth of information processing and ease of perceptual processing to facilitate learning. It also offers the possibility to communicate emotions more systematically and also to facilitate desired emotions in receivers of mobile multimedia messages, for instance. The underlying idea here is that the value of technology for a particular user is a particular user experience that may have many layers. For instance, from the point of view of a sender of an MMS-message the feeling of getting a task done is one layer. Another layer may the feeling of be getting the task done very efficiently. Yet another layer may be a feeling of accomplishing the task efficiently and enjoying doing it, lifting one´s mood. From the point of view of being a recipient of information, such as receiving an emotionally manipulated MMS-message, the choice of form of the message, or the emotionally positioned graphic template may indeed influence one´s emotion and mood in a certain manner.

Naturally, various factors influence user experiences. Detailed studies need to be done in order to understand the possibilities of varying user experiences in mobile multimedia messaging. Despite the results provided in our research, detailed guidelines on how exactly to facilitate positive emotion, for instance, are beyond the boundaries of this short article. However, it can be roughly suggested that matching the modality of information to the cognitive style of the user may provide for more efficient learning, less cognitive load and hence more pleasantness of processing the information. Similarly, certain types of personalities may have a preference for certain type of visual layout of the user interface, that, when matched with the personality of a certain user may provide for more positive emotion and mood when using the interface.

Moreover, there is a need to systematically model and classify tasks and contexts, such as

social situations and interactions and physical contexts from the point of view of influence to psychological effects of mobile messaging in real-life situations. If this can be done, one may propose Contextually Intelligent Psychological Customization Systems that would be able to customize information (media information or messaging, for instance) or other software resources (such as menus of available location-based services etc.) presented to users in a context-sensitive way.

The benefits of such Contextually Intelligent Psychological Customization may be seen from two viewpoints. First, the approach may offer psychologically informed situation models for massively computational and sensor-based approaches to context modelling to make computational inferences of the meaning of different contexts and situations easier. Second, it may offer a psychologically informed and context-based approach to present desired tools and other computational resources to the users to act creatively in a given context. In the case of mobile multimedia messaging the system could detect user being in a noisy café in a loud and intense conversation with a high level of arousal and present the emotionally adapted message accordingly in order for the psychological effect desired by the sender to be realized, trying to make the stimuli more intense and salient vs. the background interference.

Finally, it is clear the empirical validation of cognitive processing efficiency and likely emotional reactions per form factor per user profile in mobile messaging remains a laborious task let alone doing this in real-life contexts. Hence, it may be sensible to concentrate only on some key areas of application, such as facilitating selected emotions with mobile multimedia messaging systems within given contexts. Also it is evident that the collaboration of several interdisciplinary research groups may be needed to realize the emerging empiric research agenda and cross-validate the results in different cultures.

References

Billmann, D. (1998) Representations. In Bechtel, W. and Graham, G. (1998) A companion to cognitive science, 649-659. Blackwell publishers, Malden, MA.

Cuperfain, R. and Clarke, T. K. (1985) A new perspective on subliminal perception. Journal of Advertising, 14, 36-41.

Egan, D. E. (1988). Individual differences in human-computer interaction. In: M. Helander (Ed.), Handbook of Human-Computer Interaction, p. 543 – 568. Elsevier, New York.

Kallinen, K., & Ravaja, N. (in press). Emotion-related effects of speech rate and rising vs. falling background music melody during audio news: The moderating influence of personality. Personality and Individual Differences.

Kihlström, J. F., Barnhardt, T. M. and Tataryn, D. J. (1992) Implicit perception. In Bornstein, R. F. and Pittmann, T. S. (eds.) Perception without awareness. Cognitive, clinical and social perspectives, 17-54. Guilford, New York.

Krosnick, J. A. , Betz, A. L., Jussim, J. L. and Lynn, A. R. (1992) Subliminal conditioning of attitudes. Personality and Social Psychology Bulletin, 18, 152-162.

Laarni, J. (2003). Effects of color, font type and font style on user preferences. In C. Stephanidis (Ed.) Adjunct Proceedings of HCI International 2003. (Pp. 31-32). Crete University Press, Heraklion.

Laarni, J. (2002). Searching for optimal methods of presenting dynamic text on different types of screens. In: O.W. Bertelsen, S. Bödker & K. Kuutti (Eds.), Tradition and Transcendence. Proceedings of The Second Nordic Conference on Human-Computer Interaction, October 19-23, 2002, Arhus, Denmark (Pp. 217 – 220).

Laarni, J. & Kojo, I. (2001). Reading financial news from PDA and laptop displays. In: M. J. Smith & G. Salvendy (Eds.) Systems, Social and Internationalization Design Aspects of Human-Computer Interaction. Vol. 2 of Proceedings of HCI International 2001. Lawrence Erlbaum, Hillsdale, NJ. (Pp. 109 – 113.)

Laarni, J., Kojo, I. & Kärkkäinen, L. (2002). Reading and searching information on small display screens. In: D. de Waard, K. Brookhuis, J. Moraal, & A. Toffetti (Eds.), Human Factors in Transportation, Communication, Health, and the Workplace. (Pp. 505 – 516). Shake, Maastricht. (On the occasion of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting in Turin, Italy, November 2001).

Lang, A. (1990) Involuntary attention and physiological arousal evoked by structural features and mild emotion in TV commercials. Communication Research, 17 (3), 275-299.

Lang, A., Dhillon, P. and Dong, Q. (1995) Arousal, emotion and memory for television messages. Journal of Broadcasting and Electronic Media, 38, 1-15.

Lang, A., Newhagen, J. and Reeves. B. (1996) Negative video as structure: Emotion, attention, capacity and memory. Journal of Broadcasting and Electronic Media, 40, 460-477.

Ravaja, N. (2004). Effects of a small talking facial image on autonomic activity: The moderating influence of dispositional BIS and BAS sensitivities and emotions. Biological Psychology, 65, 163-183.

Ravaja, N., & Kallinen, K. (in press). Emotional effects of startling background music during reading news reports: The moderating influence of dispositional BIS and BAS sensitivities. Scandinavian Journal of Psychology.

Ravaja, N., Kallinen, K., Saari, T., & Keltikangas-Järvinen, L. (in press). Suboptimal exposure to facial expressions when viewing video messages from a small screen: Effects on emotion, attention, and memory. Journal of Experimental Psychology: Applied.

Riding, R. J. and Rayner, S. (1998) Cognitive styles and learning strategies. Understanding style differences in learning and behavior. David Fulton Publishers, London.

Riecken, D. (2000) Personalized views on personalization. Communications of the ACM, V. 43, 8, 27-28.

Saari, T. (2001) Mind-Based Media and Communications Technologies. How the Form of Information Influences Felt Meaning. Acta Universitatis Tamperensis 834. Tampere University Press, Tampere 2001.

Saari, T. (2002) Designing Mind-Based Media and Communications Technologies. Proceedings of Presence 2002 Conference, Porto, Portugal.

Saari, T. (2003a) Designing for Psychological Effects. Towards Mind-Based Media and Communications Technologies. In Harris, D., Duffy, V., Smith, M. and Stephanidis, C. (eds.) Human-Centred Computing: Cognitive, Social and Ergonomic Aspects. Volume 3 of the Proceedings of HCI International 2003, pp. 557-561.

Saari, T. (2003b) Mind-Based Media and Communications Technologies. A Framework for producing personalized psychological effects. Proceedings of Human Factors and Ergonomics 2003 -conference. 13.-17.10.2003 Denver, Colorado.

Saari, T. and Turpeinen, M. (2004) Towards Psychological Customization of Information for Individuals and Social Groups. In Karat M-C., Blom, J. and Karat. J.h. (eds.) Personalization of User Experiences for eCommerce, Kluwer.

Turpeinen, M. (2000) Customizing news content for individuals and communities. Acta Polytechnica Scandinavica. Mathematics and computing series no. 103. Helsinki University of Technology, Espoo.

Turpeinen, M. and Saari, T. (2004) System Architechture for Psychological Customization of Information. Proceedings of HICSS-37- conference, 5.-8.1. 2004, Hawaii.

Vecchi, T., Phillips, L. H. & Cornoldi, C. (2001). Individual differences in visuo-spatial working memory. In: M. Denis, R. H. Logie, C. Cornoldi, M. de Vega, & J. Engelkamp (Eds.), Imagery, language, and visuo-spatial thinking. Psychology Press, Hove.

An Analysis of Learning Effect for Paired Collaborative Learning

Kazuhiro Shin-ike
Toba National College of Maritime Technology
1-1, Ikegami, Toba, Mie, 517-8501, Japan
shin-ike@toba-cmt.ac.jp

Abstract

In paired collaborative learning some pairs improve the learning effect and others do not. In this paper, we examine factors to improve the learning effect in collaborative learning for college students who take classes of an experiment. Therefore a system for collecting some data in collaborative learning (called chat system) and a system for extracting important utterance sentences are applied to an experiment for the collaborative learning. First each student solves some problems individually. Secondly the students are divided into pair groups in order to solve some learning problems collaboratively. Thirdly, the utterance sentences of the students and their answers of the learning problems are collected by using the chat system. Then important utterance sentences for solving the learning problems are extracted from these utterance sentences by using the system to extract important utterance sentences. As a result this analysis provides which pairs can improve the learning effect. This research could be very useful to analyze the learning process in paired collaborative learning at school.

Keywords: aptitude test, learning effect, collaborative learning, chat, utterance sentences

1 Introduction

There is a proverb that two heads are better than one. According to this proverb, it may be expected that the learning effect of paired collaborative learning becomes higher than that of individual learning. T. Okada and H. A. Simon analyzed the learning effect for an individual learning and a paired collaborative learning (Okada and Simon, 1997). In their research, the learning effect for the paired collaborative learning is higher than that of the individual learning with good possibility. However, the learning effect is not always improved in collaborative learning (Kameda, 1997). Moreover, paired collaborative learning is sometimes less effective than individual learning (Shin-ike et al., 2001). Such a phenomenon is also observed in the experiment which high school students participated in (Shin-ike et al., 2002a). In our earlier paper, aptitude ability and cooperative ability of pairs of students were investigated in collaborative learning (Nakamine et al., 2002). It was found from this investigation that cooperative ability of each student was important to improve the learning effect. The method to determine of the optimal paired collaborative learning group is proposed by analyzing this investigation (Shin-ike et al., 2002b).

In this paper, paired collaborative learning is adopted in college to improve the learning effect for college students. Synthetic Personality Inventory (SPI) test is used in order to measure aptitude abilities for each student (Kaga, 2000). Ten kinds of learning problems are chosen for individual learning and paired collaborative learning. The collecting system of pairs' utterance sentences is constructed by using java software. This system enables to perform a chat communication for the learning problems by using two computers. The utterance sentences of a pair of students are recorded on the computers as a file. Moreover, important utterance sentences for solving the learning problem are extracted by using the important utterance sentences extracting system. As a result this analysis provides which kinds of pairs can improve the learning effect.

Table 1: The classification of SPI test

Inspection classification	Inspection name	Kind of inspection	The number of problems	Time limit (minutes)
Capability inspection	Inspection 1	Linguistic competence	40	30
	Inspection 2	Non-linguistic competence	30	40
Personality test	Inspection 3 and 4	Character and volition	500	60

2 Aptitude ability test and learning problem

2.1 Aptitude ability test

Synthetic Personality Inventory (SPI) test is used for inspecting the aptitude ability. SPI test is most utilized as the employment examination of private enterprises in Japan. It includes both the capability inspection and the personality test and it was developed by one of Japanese companies in 1974. The SPI test is a comprehensive aptitude inspection, and individual capability and character are measured and evaluated. It is said that 10,000 top-rated companies in Japan mainly adopt it now, and it is utilized as important data for employing people.

Table 1 shows the contents of SPI test. SPI test can tell where a person has more than a fixed level of the linguistic competence (language capability) and the non-linguistic competence (mathematical capability) in capability inspection. The score of this test is utilized as a cut-off point to select applicants who pass the first examination. The personality test is used in order to know an applicant's character and volition. Individual character is divided into four sides; emotional side, active side, volitional side and characteristic side. This result is used as important basic data of an interview for the second examination.

All the questions of the test have some alternatives of answers and the thinking time per question is quiet short. Inspection items are shown below.

1. Inspection 1 (Linguistic competence (language ability))
   It is made of linguistic questions and linguistic competence is measured. For example, there are the ability to understand the meaning of the words and phrases, reading ability, writing ability, etc.

2. Inspection 2 (Non-linguistic competence (mathematical ability))
   The questions of mathematics and science are applied. The level of these questions is for the students from the seventh grade to the tenth grade. Each question of the graphs of functions, equations, physics, etc. is selected from any fields.

3. Inspection 3 and 4 (Personality test)

   (a) Emotion-side
   This inspection is to measure emotive stability and the adaptability in an organization.

   (b) Active-side
   This inspection is to the side where the character tends to appear.

   (c) Volitional side
   This inspection is to measure the level of motivation, volition and vitality.

   (d) Type of character
   This inspection is to measure the interest, concern, and judgment ability that the examinee's has.

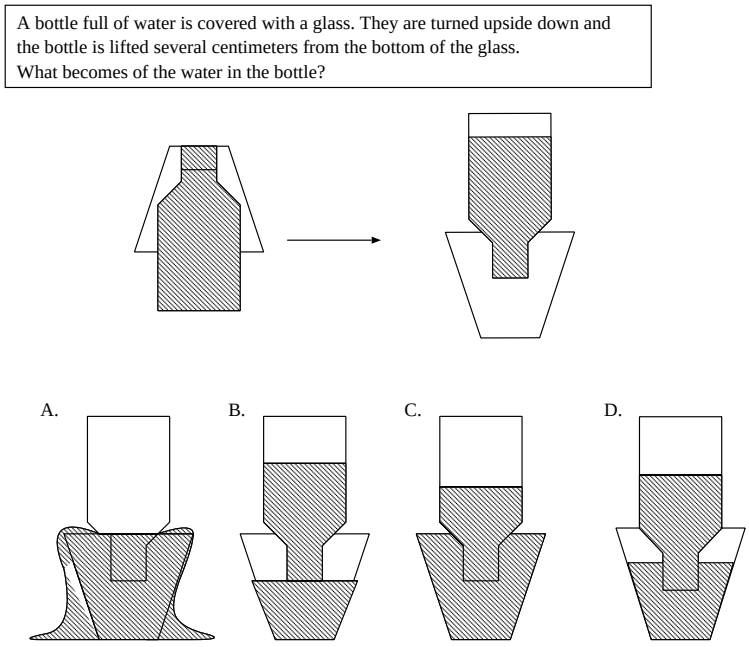


Figure 1: The learning problem example 1

2.2 Learning problem

The learning problems for the individual learning and the collaborative learning are selected from the phenomena in a daily life. Figure 1 shows one of the learning problems. This figure shows that a bottle full of water is covered with a glass. Next, they are turned upside down and then the bottle is lifted several centimeters from the bottom of the glass. Then the water in the bottle flows into the glass. If the bottle keeps at the same position, what becomes of the water in the bottle? This is one of the learning problems. In this problem, four kinds of answers are prepared as shown in figure 1.

The correct answer of this problem is B. This problem is related to the atmospheric pressure. There is air pressure at the water surface of the glass, and this air pressure works as power that pushes the water into the bottle. On the other hand, air and water are in the bottle. The resultant of the weight of the water and the pressure of the air works as pressure that pushes out water from the bottle. When the pressure to push in the water equals to the pressure to push out it, the flow of water stops.

The number of the learning problems is ten and the alternatives of the answers are prepared.

3 A system for collecting some kinds of data in collaborative learning

3.1 A system for collecting utterance sentences by using a chat

A chat system is used for collecting the utterance sentences of collaborative learning, which are the contents of the utterance sentences of a pair of students for the learning problem, the answers of the problem they select and the time when they are learning. A network system is necessary to collect these three kinds of data.

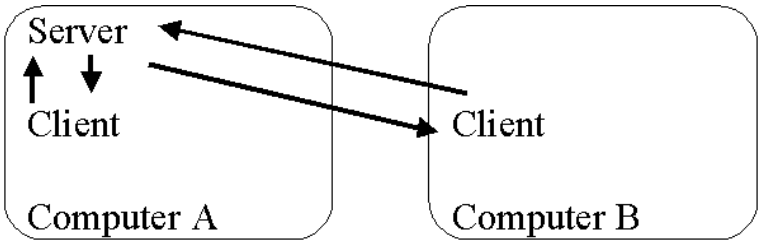


Figure 2: Chat system

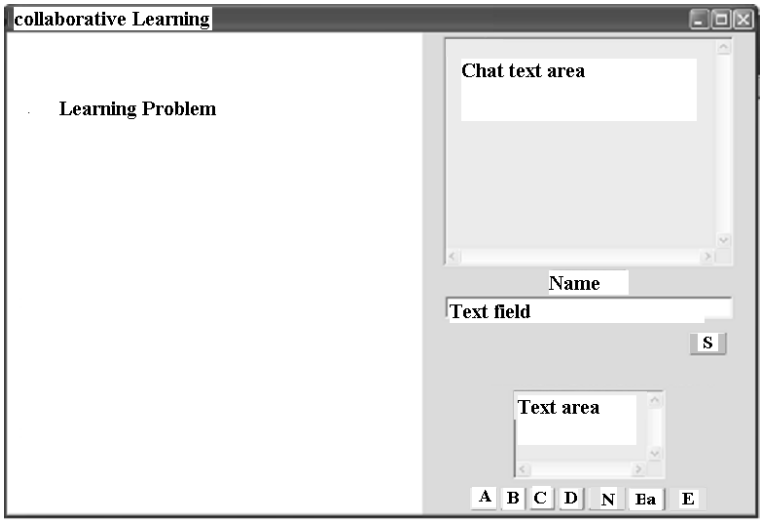


Figure 3: The layout of chat system

The computers connected to the network exchange data by using protocol. In this research, TCP (Transmission Control Protocol) is used to send and receive data and to decide the order of data. In the case of communication between computers, all-important data are stored in a server and data of the server are provided to clients on the network. For example, a Web server provides the specified page to clients.

In the server program and the client program that carries out TCP communication, the server prepares a socket and waits for a connection request from a client. If a client makes a connection request, the request will be received and the server will prepare a new socket. At this time, the original socket waits for a connection request from another clients. Thus, the client communicates with a server through a socket. Figure 2 shows that two students communicate with two computers. Computer A supplies the program of both a server and a client in the chat system. On the other hand, computer B supplies the program of the client. The client enables each computer to communicate mutually through the server in Computer A. Although Figure 2 shows the communication by using two computers, three or more computers can be used, too.

Figure 3 shows the interface layout of this system. When this system is executed, the explanation of the usage is displayed on the screen. If the text that one of the pair of students wants to transmit to a partner is inputted into the text field shown in this figure and a transmitting button is clicked, the name of the student who clicks the button and the transmitted contents are displayed on text field shown in this figure.

The problem number 1 will be displayed if N button is clicked (refer to Figure 3). If one of the answers buttons A, B, C and D for this problem shown in this figure is clicked, the number of this problem and the answer are displayed on text area at the same time. If the answer selected by a

pair of students is correct, the description of this problem is displayed and if the O.K. button is clicked, the following problem will be displayed.

If this answer is incorrect, the same problem will be displayed again. A pair of students repeats this work until it answers ten kinds of learning problems. When they finish answering problems, the collaborative learning is completed by clicking E button shown in figure 3.

When E button is clicked, the contents of utterance sentences for a pair of students are saved in a chat.txt file, and the answers of the learning problems are saved in a data.txt file. These files are saved in a folder of the program of a chat system. Moreover, the learning time is displayed on an MS-DOS prompt.

3.2 Programming with Java

In the case of the experiment for the collaborative learning, it is necessary to collect contents of utterance sentences by students and their answers to the learning problems. A system to collect them is constructed with java software. In this system the learning problems and some choices about the answers to the learning problems are displayed and the answers of students are recorded. Moreover, it is possible to perform a chat communication on the same screen as the learning problems with two computers.

Java is the object-oriented language that Sun Microsystems developed in 1995. The reason for using Java is as follows. First, it is easy to create window components with java software. In this research, students need to solve the learning problems by using easy operation on a computer screen. Java software can create the application (called window component) that shows various kinds of outcomes on a computer screen by pushing a button or inputting some characters with keyboard by using the simple program.

Secondly, the communication on a network can be made smoothly. It is important to communicate on the network because in this research it is necessary to record the utterance sentences of the students for the collaborative learning. Moreover, when the students use this chat system, two or more clients can be processed with a thread at the same time. The thread means a flow of a series of processing in an execution program. It is possible for java software to program by using multi-thread. In java software, it is possible to perform two or more operations through a network at once.

Finally, it is the feature of java software that it does not depend on the platform. Since it does not matter the types of computers or the operating systems, this system is actually used at schools etc.

4 A system for extracting important sentences

In this research we consider each student's thinking process and interaction when students solve a learning problem together, so utterance sentences of each student are collected by the chat system. This system can store these utterance sentences as a text (called chat text). The words that carry out the role of the keywords in a text are called important words. Important words are fundamentally extracted by the frequency of words. Furthermore, the important sentences carrying out the important role in order to solve a problem are extracted by using the important words.

The high-frequency words can be regarded as a big feature in the chat text. However, some low-frequency words carry out an important role for problem solution in the chat text. In this research, in order to extract the words showing the feature of the chat text, a technique shown below is applied.

4.1 Extraction of high-frequency words

A chat sentence is defined by s and the set is S ($s \in S$). Evaluation value $r_w(w_i)$ of each word w_i is given by the formula (1) as a frequency rate in S.

$$r_w(w_i) = \frac{frequency(w_i)}{\sum_i frequency(w_i)} \tag{1}$$

where $frequency(w_i)$ is the frequency of appearance of word w_i. In many cases, the word used two or more times is carrying out the important role in the text. About 20 pieces of words high evaluation value is determined on candidate g of important words. The set is taken as G ($g \in G$).

4.2 Extraction of important words

The word that carries out an important role even if frequency rate is low may appear in the text containing important words. If the same word is used for different sentences, the importance of the word in each sentence differs. The dignity of the word is performed as follows.

Dignity w_w of the word w_i in sentence $s_j \in S$ is expressed with the number of words determined with section 4.1. $G \cap s_j$ is shown the number of words contained in both s_j and G. If $n(G \cap s_j)$ is the number of words of $G \cap s_j$, w_w is given by the following formula.

$$w_w(w_i, s_j) = n(G \cap s_j) \qquad (w_i \in s_j, s_j \in S) \tag{2}$$

$v_w(w_i)$ shows total of word w_i calculated about each text using the formula (2). This total is called evaluation value of word w_i, and is expressed with the following formula.

$$v_w(w_i) = \sum_{s_k \in S} w_w(w_i, s_k) \tag{3}$$

A maximum of h words are chosen as order with the large evaluation value. These words are called important words. Set H of important words is given by the following formula.

$$\begin{aligned} H \quad = \quad & \{w_{i_l};\ l = 1, 2, \cdots, h \\ & \mid v_w(w_{i_1}) \geq v_w(w_{i_2}) \geq \cdots \geq \\ & v_w(w_{i_h})\} \end{aligned} \tag{4}$$

4.3 Extraction of important sentences

The evaluation value of each sentence is determined by using the important word calculated by the formula (4). The evaluation value $v_s(s_k)$ of a sentence s_k is given by the following formula.

$$v_s(s_k) = \sum_{w_{i_l} \in s_k \cap H} v_w(w_{i_l}) \tag{5}$$

The number of important sentences is determined by the value $v_s(s_k)$ calculated by the formula (5).

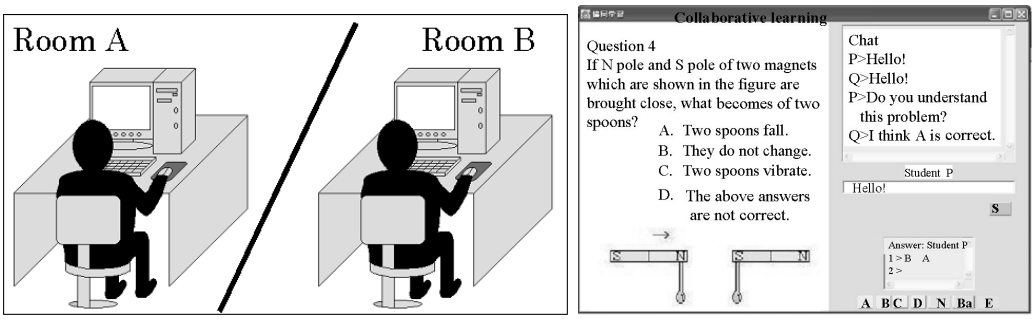


Figure 4: The experiment by using chat system

5 Experiment

5.1 Paired collaborative learning

Eighteen students participate in the experiment for the collaborative learning. They are twelve male sophomoric students and six female sophomoric students at college. As they have been studying subjects for five years in the same classroom with the same curricula, they are able to talk their opinion each other freely. They are assigned to nine pairs by experimenters. Six pairs are mixed gender students and three pairs are the same gender students. The experimenters verify each student's learning situation and classify the students into the pair whose learning effect can be improved, and the pair whose learning effect cannot be improved. The experimenters predict that the learning effect of five pairs is improved, and that of four pairs goes down.

First of all, each student takes SPI test for investigating the individual personality and solves ten kinds of learning problems. The correctness rates of the learning problems in the case of individual learning are investigated. The problems are the same as those of the collaborative learning.

Next, pairs of students take an inspection of the collaborative learning with the chat system. Left hand side of figure 4 shows the scene of the experiment to get the answers of learning problems, the contents of utterance sentences of the pair of students and the time for answering to the problems. The pair of students communicates with each other by inputting words with keyboard.

Right hand side of figure 4 shows the screen of the computer of student P who is solving the problem with student Q in collaborative learning. The problem is displayed on the left-hand side of the screen. The utterance sentences which two students exchange are displayed on the top of the right-hand side of the screen by letter. The name of the student who is using the computer is displayed in the center of the right-hand side of the screen. The text, which student P transmits to student Q, is inputted into the text field under the name of student P. The answers of the problems that two students choose are displayed on the text area at the bottom of the right-hand side of the screen. It is found from the text area in this figure that two students chose B for the first time as the answer of the problem 1 and chose A for the second time. As a matter of fact B is a wrong answer and A is a correct answer.

Table 2 shows the correctness rates for ten kinds of learning problems in individual learning and collaborative learning. In this table, it is found that the correctness rates of two pairs (pair number 1 and 2) in collaborative learning are lower than that of each of those students in individual learning. Pair number 1 that consists of male gender students solves the learning problem without discussing with each other. Pair number 2 consists of mixed gender students. It is found from the experiment that the cause that the correctness rate in collaborative learning of this pair is lower than that of each student in individual learning is not the difference of gender, but their aptitude ability and personality: It is found from SPI test that one student of pair number 2 is too assertive and egotistic. The other student is not so active and does not express his or her opinion very much. So, this pair rarely talks about how to solve the learning problem each other.

Table 2: The correctness rates of individual learning and paired collaborative learning

Pair No.	1	2	3	4	5	6	7	8	9
Student (i)	0.2	0.6	0.4	0.5	0.3	0.7	0.6	0.3	0.2
Student (j)	0.6	0.6	0.2	0.4	0.5	0.3	0.6	0.6	0.6
Pair (i, j)	0.0	0.5	0.2	0.4	0.5	0.5	0.6	0.6	0.7

Table 3: Examples of answer history of each student for the learning problems

Problem No.		1	2	3	4	5	6	7	8	9	10
Pair (i, j)	i	C	B	B	C, A	A	C	C, B	C, B	A	C
	j	C	B	B	C, A	A	C	C, B	C, B	A	C
Pair (k, l)	k	C	C, D, B	D, B	B, D, C, A	C, B, D, A	B, C	A, B	A, C, B	A	C
	l	D, C	C, B	B	D, C, B, A	A	A, C	B	C, B	A	C

Each of the correctness rates of six pairs of students (pair number 3, 4, 5, 6, 7 and 8) in collaborative learning is between the correctness rate of student (i) and that of student (j) in the individual learning. Pair number 3, 4, 5, 7 and 8 consist of mixed gender students and pair number 6 consists of male gender students. Pair number 4 is analyzed as an example of these six pairs. It is found from SPI test that the student (i) is a female of progressive, careless and active personality and the student (j) is a male of careful personality. Consequently, when the student (j) explains the meaning of the learning problem to the student (i), the interaction of these two students is not so good because of carelessness personality of the student (i) and as result the learning effect in collaborative learning does not improve.

In the case of pair number 9 the correctness rate in collaborative learning is higher than that of each student in individual learning. This pair consists of male gender students. One of this pair of students is a male of considerate and permissive personality and the other is a male of communicative and active personality. Consequently, this pair is able to discuss the learning problem usefully and the learning effect in collaborative learning is better than that of each student in individual learning.

Table 3 shows two examples of answer history obtained by using the chat system. In this table, A, B, C and D are the choice of the answers of the learning problems. It is shown that both students of Pair (i, j) (pair number 9 shown in Table 2) select C as the correct answer of the problem 1 for the first time, select C as the correct answer of the problem 4 for the first time, and select A as the correct answer of this problem for the second time. On the other hand, each student of Pair (k, l) (pair number 1 shown in Table 2) makes a choice without discussing the learning problems. Especially, student k makes a choice 3 times or more in four kinds of learning problems. It is found that the pair of students i and j learns in collaboration with each other but the pair of students k and l does not.

5.2 Utterance sentences of a pair of students

Table 4 shows utterance sentences of a pair of students P and Q (pair number 9 shown in Table 2) solving the learning problems shown in figure 5 by using the chat system. They argue about the chemical character of water eagerly in order to solve the problem. Their utterance sentences include the open question containing the keywords in the learning problem. Open question means urging a partner to talk freely such as "What do you think? " and "Why? ". Moreover one of the students talks adding polite explanation of the learning problem. And the other student asks some questions and answers simply like "Yes" or "No". It seems that each student not only says his opinion, but also explains the learning problem or listens to the explanation of the partner. As a result the learning effect in collaborative learning becomes higher than that in individual learning.

Table 4: An example of utterance sentences saved by using chat system

Name	Utterance sentences
P	Hello!
Q	Hello! Now! Do you know the correct answer of this problem?
P	I think it is D.
Q	I think it is C.
P	Why do you think the correct answer is C?
Q	Ice is lighter than water, isn't it?
	They differ in density.
P	T upper of tha ice is out of water in a glass. If ice melts
	away, I think that water overflows the glass.
Q	The weight of the ice is equal to the weight of
	the water of the portion which is underwater.
P	I think so.
Q	Since the density of ice is smaller than that of water, the ice floats on the water.
P	Great indeed.
Q	I do not understand.
P	I think that C is correct.
Q	I think so, too.
P	But is there no possibility of B?
	I do not understand in the least.
Q	Neither ice nor water goes anywhere.
P	I think that the correct answer is C. I think that the problem was not read correctly.
Q	The water surface is heaped up.
P	It was late for me to notice that.
Q	You should read the learning problem more carefully.
P	Just as you same.
Q	Then, the correct answer is C.

On the other hand, the pair of students (pair number 2 shown in Table 2) has many useless conversations and misunderstandings the learning problem. For example, if water freezes, the volume of water will decrease while melting. Consequently the learning effect in collaborative learning of this pair is lower than that of each of them in individual learning

Correctness rates of six pairs of students (pair number 3, 4, 5, 6, 7 and 8 shown in Table 2) in collaborative learning are between the correctness rates of two students in each pair in individual learning. The main features of these pairs of students are a few utterance sentences and the lack of knowledge about the state change of water. It is found from their utterance sentences that they do not think about the learning problems well and they have some dubious utterance sentences. The dubious utterance sentences between the pairs of students have a bad influence in collaborative learning. Consequently these pairs of students do not improve the learning effect in collaborative learning.

6 Important sentences for improving the learning effect

In this section, the system shown in section 4 is applied to some pairs of students. The number of words shown in table 4 is two hundred and eight. First, high frequency words are extracted from the text of table 4 by using the system shown in section 4.1. Secondly keywords for solving the learning problem shown in figure 5 are extracted from the text on the basis of the high frequency words. Finally important sentences for solving this learning problem are extracted from the text on the basis of the keywords. Table 5 shows the keywords and the important sentences extracted from the utterance sentences.

Six words "light, density, ice, overflow, weight, portion" in these eight keywords are concerned

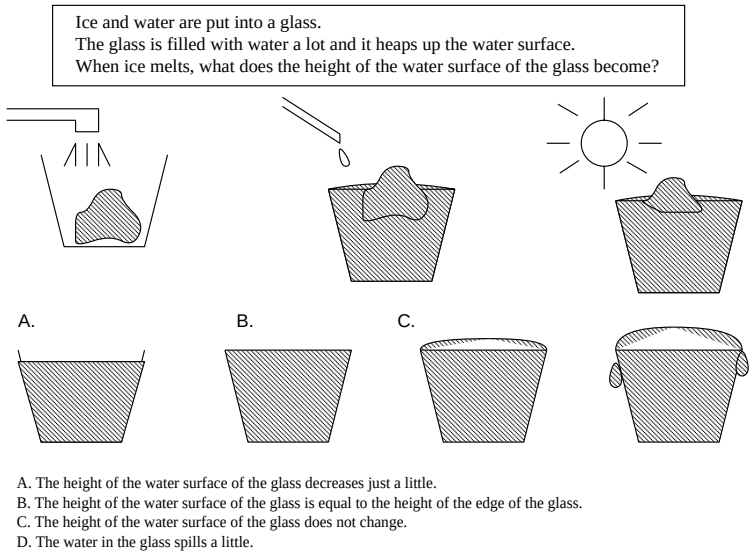


Figure 5: The learning problem for example 2

Table 5: An example of keywors and important sentences

keywords	light, density, ice, overflow, weight, portion, C, understand
Name	Important sentences
Q	Ice is lighter than water is not it. They differ in density.
P	Since there is ice on a glass, if ice melts away, I will think that water overflows from a glass.
Q	The weight of the ice is equal to the weight of the water of the portion which is underwater.
P	I think that an answer is C. I think that the problem is not read correctly.
Q	Then, the answer is C.

with the ice and the water in a glass shown in figure 5. The word "C" is the answer of the learning problem. The word "understand" is used when the students express their thought.

It is found from this table that the keywords extracted by using the proposed method are effective in solving the learning problem. It is also found from the important sentences that this pair of students has conversation with the open question containing the keyword in the learning problem. Moreover it is found that one of the pair talks adding polite explanation about the learning problem. Consequently, the correctness rate of this pair in collaborative learning is higher than that of each of them in individual learning. Therefore, this proposed system shown in section 4 is effective to analyze the interaction of pairs of students in this collaborative learning.

7 Conclusion

In this research, experiments of individual learning and paired collaborative learning are adopted to find the factors to improve the learning effect in paired collaborative learning. The chat system is applied to collect the utterance sentences of pairs of students and the answer history of them. The data obtained from this experiment is analyzed. It is found that the pairs of students who get high correctness rates can make the sentences containing their opinions, and each student of the pair can fill the role. This chat system is applicable to collecting the utterance sentences in

collaborative learning by using the Web page.

Next, the extracting system of important utterance sentences is proposed. This system is applied to the paired collaborative learning. It is found that the keywords obtained by the utterance sentences show the feature of the learning problem as shown in figure 5. The important utterance sentences can be extracted from the utterance sentences of a pair of students by using the keywords. It is found from these important utterance sentences that the learning situation and the role of the pair can be understood in the case of the paired collaborative learning. The time to extract important utterance sentences is reduced by using this system. It takes shorter time than using the conventional system.

These two systems in this research could be very useful to analyze the learning process in paired collaborative learning at school. In this paper, there are not many experimental pairs, so the aim of the future is to increase experimental pairs and is to analyze the learning effect in paired collaborative learning in more detail.

References

Kaga H. (2000). Practice Exercise of SPI Test, *Nagaoka Syoten*. (in Japanese)

Kameda T. (1997). Comparison of Individuals and Groups in the Rational Solution of Complex Problems, *American Journal of Psychology*, Vol. 44, 491-504.

Nakamine H., K. Shin-ike & N. Sannomiya (2002). Analysis of Combination Types in Paired Collaborative Learning by Using a Self-Organizing Neural Network, *Proc. of the 6th World Multiconference on Systemics, Cybernetics and Informatics*, 202-207.

Okada T. and H. A. Simon (1997). Collaborative Discovery in a Scientific Domain, *Cognitive Science*, Vol. 21, No. 2, 109-146.

Shin-ike K., H. Nakamine & J. Kirita (2001). A Neural Network for Learning Effect Prediction in Collaborative Circuit Board Design Learning, *Journal of the Japanese Society of Technology Education*, Vol.43, No. 2, 85-92. (in Japanese)

Shin-ike K., H. Nakamine, & N. Sannomiya (2002a). A Method for Development of Paired Collaborative Learning in a Technical High School, *Proc. of the 2002 American Society for Engineering Education Annual Conference*, Session No. 2793.

Shin-ike K., H. Nakamine and N. Sannomiya (2002b). A Method for Determing Pairs of Students in Collaborative Learning by Using Neural Networks and Local Search Method, *Trans. of the Society of Instrument and Control Engineers*, Vol. 38, No.5, 477-484. (in Japanese)

Selfish Agents and the Butler Paradigm

Kees Tuinenbreijer,
Twente University, Enschede, the Netherlands
and Philips Electronics, Eindhoven, the Netherlands,
kees.tuinenbreijer@philips.com

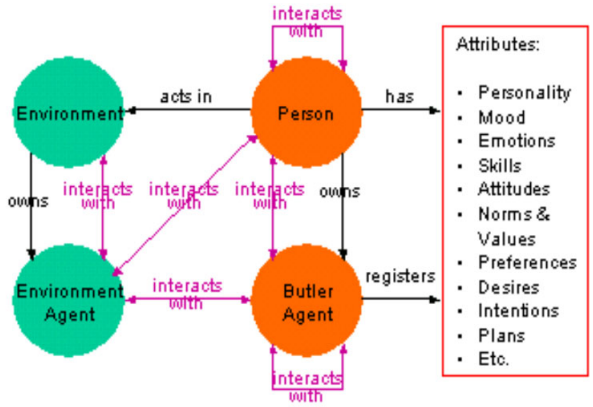
Abstract

The University of Twente does research on the Butler paradigm. The Butler paradigm comprehends *social* butler-agents in intelligent environments (home and public). Butler-agents act on behalf of its owners and are aware of owner's attributes such as personality, mood, emotions, goals, attitudes, values, plans and more. Butler-agents know about its owners in restricted terms; restricted by the simplicity of the used models.
In the context of the Butler paradigm, this paper highlights social characteristics of butler-agents and pays attention to the need of an adequate 'individual-society' social balance required to ensure that butler-agents are able to function in an effective manner.

Keywords: Social Interaction, Selfish Agents, Multi Agent Society.

1 Introduction of the Butler Paradigm

The main entities and their relations in the Butler Paradigm are:



In the context of this paper, the following is relevant:

- A *Butler-agent* acts on behalf of its owner (*Person*) and is aware of owner's attributes and *Environment*'s (society) norms.
- In this *Environment*, *Butler-agents* support valued social relationships and have a clearly defined social role.
- *Butler-agent*s are (semi-) autonomous agents and maintain awareness of each other activities and intentions.
- *Butler-agents* are consistency, reactive and responsive, situated and resource-bounded.

- The levels of responsiveness, intrusiveness, reactiveness and proactiveness of *Butler-agents* are adjustable by its owners.

2 The 'Mrs. Smith' scenario

To make the Butler paradigm more concrete and illustrative, a scenario is added to this paper. It is called the 'Mrs. Smith' scenario:

> *Mrs. Smith is a busy person. The butler of Mrs. Smith (an agent) supports Mrs. Smith and is always with her. Mrs. Smith's butler has the capability to react on Mrs. Smith's physiological attributes, such as blood pressure, temperature and more. Mrs. Smith's butler is also aware of other personal attributes, such as personality, mood, emotions, motivation, goals etc. Mrs. Smith's butler is a learning agent: In the beginning, Mrs. Smith has partly 'programmed' her butler, but now the butler learns about Mrs. Smiths from her physical condition and her interactions with her environment. Mrs. Smith's butler is an (pro-) acting agent and offers communication, processing and decision-making functionality. This morning, Mrs. Smith is not feeling well. Her temperature is not normal. Mrs. Smith's butler is worried and he initiates a small talk with Mrs. Smith. Mrs. Smith acknowledges her feelings of little sickness and she asks her butler to arrange a doctor-visit. After studying Mrs. Smith's agenda and consulting the butler of her family doctor (agent), Mrs. Smith's butler suggests a visit in the afternoon. (In the morning is not possible because of an important meeting with Mrs. Smith's development staff). Mrs. Smith agrees with the visit and her butler goes into action to manage the necessary rearrangements in Mrs. Smith's and the doctor's agenda. The socially sophisticated butler of the doctor has explicit knowledge about the enacted norms in the multi-agent environment of the Butler paradigm. Moreover, the butler of the doctor has experiences and knowledge of the interests and behavior of (the butler of) Mrs. Smiths. Based on this, the butler of the doctor is able to react on a social way to requests of the butler of Mrs. Smith.*

3 Social interaction

There are two fundamental questions concerning social interaction:

- Sociality problem: why does an autonomous agent enter into social interactions?
- Adoption problem: how does an agent get his problem to become social, i.e., get it adopted by other agents?

The answer to the first question is social dependence: an agent depends on another one if this latter can facilitate the former to achieve one of his goals. The answer to the second one is social power: is an agent depends on the other one, than the second agent has power on the first one. Basically, a 'lack of power' concerning their own goals make agents dependent and their 'power to influence' someone else leads to goal adoption (Castelfranchi, 1990). In the study on the Butler paradigm, Castelfranchi's model of social power and dependence is adopted (Castelfranchi et al., 1992; Conte and Castelfranchi, 1995).

4 In a Social Environment

A key hypothesis of the study on the Butler paradigm and other studies (Kalenka 2001) is that the best means of building socially coherent multi-agent systems is to endow the Butler-agents with social awareness to strike a balance between own needs and those of the environment. The best way of attaining such a balance is to develop Butler-agents which retain their local autonomy, are able to benefit from interactions with other agents, but which are also willing to provide some resources for the benefit of the overall system some of the time.

In the study on the Butler paradigm, the Principle of Social Rationality (Jennings and Campos 1997) is explored: '*If a member of a responsible society can perform an action the joint benefit of which is greater than its joint loss, then it may select that action'.*

5 In a Open Society

In general, in open societies, in which agents of any type can join at any time or change their behavior from time to time, agents do not have knowledge about the behavior and the interests of the other agents. This means that a socially responsible agent does not know if the agent it is interacting with also has (same) social interests and it is not guaranteed that all actions and resources needed to achieve certain goals are available.

An important question in an open society is how to handle and how to protect against the freeloading phenomena: social responsible agents sometimes perform actions for the greater good, which they would not have performed in case of a purely self-interested stance. Such behavior may stimulate **selfish agents** to exploit such social responsible agents without producing an adequate overall benefit for the society. The butler-agents of the doctor and Mrs. Smith are most of the time, more or less, socially responsible agents that retain their autonomy but also draw from, and provide resources to, the larger society. In general, the butler-agents attempt to balance the needs of their owners with those of the overall society. However, it is an open society and not all agents in the Butler paradigm are always socially responsible agents. An effective strategy is required to protect agents against exploitative behavior such as freeloading.

In the study on the Butler paradigm and other studies (Kalenka 2001), point of departure is the assumption that, in a *like-minded* society, the more the experiences and the knowledge about the interests and behavior of other butler-agents are taken into account; the better is the protection against exploitation. An example: the butler-agent of the doctor has to distinguish hypochondriac-like and more serious patients. In the course of time, the butler-agent of the doctor will be well informed of behavior of Mrs. Smith. Based on experiences and knowledge about the behavior of Mrs. Smith, the doctor-agent is able to judge the urgency and reasonableness of the request of the Mrs. Smiths butler-agent.

References

Castelfranchi C. (1990). "Social Power: A Point Missed in Multi-Agent DAI and HCI" in Decentralized AI (eds. Y. Demazeau and J. P. Müller), Elsevier Science Publishers B.V., North Holland.

Castelfranchi C., M. Miceli, and A. Cesta (1992). "Dependence Relations among Autonomous Agents", in: Werner and Demazeau: (eds.): Decentralized Artificial Intelligence 3, Proceedings of the 3rd European Workshop on Modelling Autonomous Agents in a Multi-Agent World, Elsevier, Amsterdam, NL.

Conte, Rosaria and Christiano Castelfranchi (1995). "Cognitive and Social Action", UCL Press, University College London.

Jennings, N.R. and Campos J. R. (1997). "Towards a Social Level Characterisation of Socially Responsible Agents", IEE Proceedings on Software Engineering, 144 (1), pp. 11-25.

Kalenka, Susan (2001). "Modelling Social Interaction Attitudes in Multi-Agent Systems", Thesis submitted for the degree of Doctor of Philosophy in the Faculty of Engineering. Queen Mary, University of London.

Designing Conversational Agents: A Psychological Approach to Aid User Comprehension

Koji Yamashita,
Social Interaction Group,
National Institute of Information Communications and Technology
3-5 Hikaridai, Seika, Kyoto 619-0289, Japan
koji@nict.go.jp

Hidekazu Kubota and Toyoaki Nishida
School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan
{kubota|nishida}@kc.t.u-tokyo.ac.jp

Abstract

We have developed a broadcasting agent system, POC caster, which generates understandable conversational representation from text-based documents. POC caster circulates the opinions of community members by using conversational representation in a broadcasting system on the Internet. We evaluated its transformation rules in two experiments. In Experiment 1, we examined our transformation rules for conversational representation in relation to sentence length. Twenty-four participants listened to two types of sentence (long sentences and short sentences) with conversational representation or with simple speech. In Experiment 2, we investigated the relationship between conversational representation and the user's knowledge level. Forty-two participants (21 with a high knowledge level and 21 with a low knowledge level) were selected for a knowledge task and listened to two kinds of sentence (sentences about a well known topic or those about an unfamiliar topic). Our results indicate that the conversational representation aided comprehension, especially for long sentences and when users had little knowledge about the topic. We explore possible explanations and implications of these results with regard to human cognition and text comprehension.

Keywords: Conversational Agents, Conversational Representation, User Comprehension, Psychological Factorial Experiment

1 Introduction

Agent technology is flourishing, and various agents have appeared on the scene. For example, Rea (Real Estate Agent) is a sales agent for real estate (Cassel, Bickmore, Campbell, Vilhjalmsson, & Yan, 2000), and STEVE is a pedagogical agent for procedural training (Rickel & Johnson, 1998). The word agent is commonly defined as meaning autonomous software that interacts intellectually with users or other agents. As such, it acts as a mediator between users and information. Conversational representation should be an effective interface for an ideal mediator because it is primitive and familiar to users and does not require particular skills (such as keyboard typing).

Conversational representation for software agents can be used in two ways. One is for information provided from agents to users, and the other is for information exchanged between users and agents. In the first case, the main goal is that agents provide information to users in an easy-to-understand way. Using conversational agents in a broadcasting system is a typical case. Agents, for example, can read texts aloud using speech synthesis technology. This is one-way communication because there is no continuous interaction between users and agents. In the latter case, agents provide information through interactions with users, interactively confirming the user's intention

and comprehension. This is a bidirectional approach, and it may be the ideal for many situations. Neither approach, however, has yielded satisfactory systems because of problems with speech synthesis (regarding both voice recognition and natural voice generation) as well as difficulties when the system cannot recognize the user's intention or generate a proper conversational representation. Overcoming these problems is crucial, especially in the latter case, because they block the natural flow of our thinking and communication. In this study, we have focused on information provision by conversational agents in a broadcasting system (i.e., one-way communication) since it should entail fewer problems than a bi-directional approach.

Information provided using the forms of everyday conversation is often found in magazines and newspapers, TV news programs, etc. Using a conversational representation allows us to picture a scene differently from when a single speaker reads a text. A listener's comprehension is thought to be facilitated by conversational representation because inserted pauses, the speaker's rhythm, and topic control by two or more speakers generate cues which help the listener understand what is being said. More specifically, our comprehension of a topic can be deepened if we ask questions and discuss it with others. Is a listener's comprehension of a topic similarly affected when agents use conversational representation in providing information? If so, what kind of presentation is most effective?

In this paper, we report on two evaluative experiments with a psychological factorial design. The experiments used a conversational agent (POC caster) that we have developed (Kubota, Yamashita, Fukuhara & Nishida, 2002). POC caster is a conversational agent using a method for transforming a text into a conversational representation. In the next section, we provide a system summary of POC caster and our transformation rules. In Experiment 1, we evaluated our transformation rules for conversational representation in relation to sentence length. More specifically, we manipulated two factors; sentence length and representation form. In Experiment 2, we examined the relationship between conversational representation and knowledge level. In the final section, we conclude this paper by exploring possible explanations and implications of the experimental results.

2 Our approach to conversational agents

We have already described POC caster, a broadcasting system conversational agent for transforming a text into a conversational representation (Kubota et al., 2002). POC caster operates on the Public Opinion Channel (POC; Azechi, Fujihara, Sumi, Hirata, Yano, & Nishida, 2000), which is an automatic interactive broadcasting system to support knowledge creation and facilitate knowledge circulation in communities on the Internet. By 'ecommunity', we mean a group of people loosely connected through shared interests or environments; for example, a group of people with similar interests, hobbies, or goals. The POC gathers the opinions of community members, edits broadcasting programs based on the collected opinions, and then broadcasts them as TV or radio programs. The POC elicits knowledge from the community and facilitates its circulation by automating this cycle. The POC system consists of the POC Server (Fukuhara, Chikama, & Nishida, 2003) and a number of POC clients (for example, POC Communicator and POC caster in Figure 1), which are connected by a network.

In the POC system, information posted by members is stored in the form of a knowledge card (POC card), which consists of a title, a channel name, a file name, the date posted, an IP address, a still picture, and text (approximately 100 words of plain text). Opinions from community members may include comments on, for example, local news, local folklore, or trivial information discovered in daily life. We anticipate that using a "small talk"style (a small information unit) will encourage community members to post their opinions and messages, because many people may find it difficult to put up contents for posting. When a POC card is broadcast to the community, it is called a message and consists of a title, text, and a still picture; that is, when the message is broadcast, the POC does not provide any information about the author. The system is designed to exclude information about the author because social psychologists believe that anonymous communication facilitates effective discussions in communities (Azechi et al., 2000). Anonymous posting seems to encourage community members to post information without hesitation, which

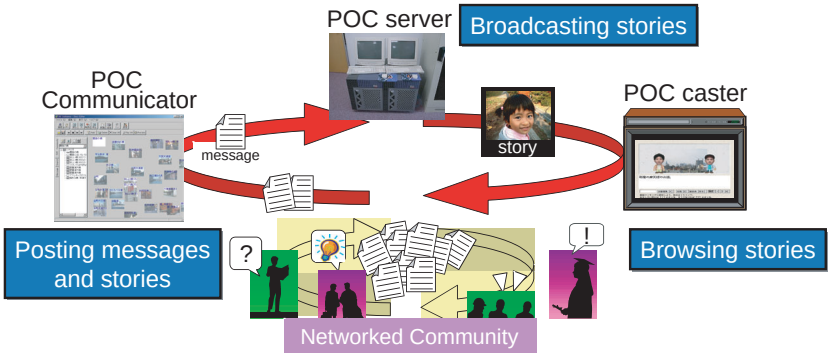


Figure 1: Public Opinion Channel (POC) Concept

helps increase the amount of message posting. POC information may also be posted in the form of a story consisting of a collection of messages. For example, messages can be combined into a story much like a picture-card show. Each community member can reuse existing messages on the POC Server as part of her/his story without needing the consent of any member of the POC community. Members can thus develop shared stories by collaboratively editing and constructing them. We believe that reusing the messages of others in this way may also lower the threshold for posting behavior. Messages and stories stored in the POC Server are always displayed circularly in the main window so that members can access them at any time. For example, when a member is creating a story, seeing a message that happens to be displayed in the main window may spark a new idea for the story composition. This function enables a "blending" of messages and stories within the community. In a nutshell, the main characteristics of the POC are (1) "small talk", (2) anonymous posting, (3) collaborative story building, and (4) circular display of stored knowledge.

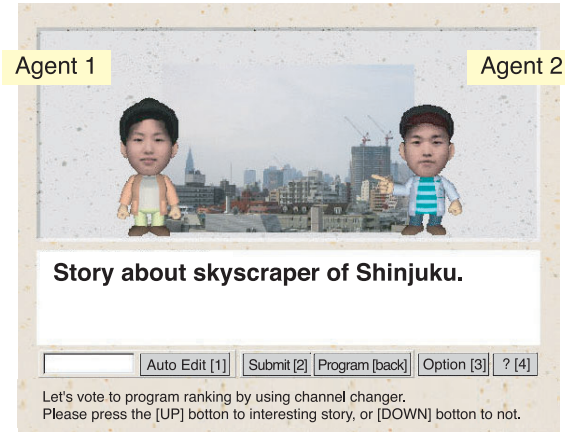


Figure 2: POC caster Screenshot

POC caster is an information-providing agent operating on the POC system. It presents information to community members by transforming opinions from other community members into a conversation between two agents. First, a user inputs a keyword to indicate interest in a certain topic. POC caster sends an inquiry to the POC Server, where opinions submitted from the community members are stored, and receives in return a set of opinions related to the keyword. The opinion set is arranged in a manner the user specifies: chronological order, reverse chronological order, or at random. POC caster transforms each opinion from plain text into a conversational representation, and broadcasts it to users by using speech synthesis with captions.

Table 1: POC caster Transformation Rules

Rule 1:	*Presentation of the context*	
	Objective	To facilitate listener's understanding of the context by inserting words
	End of sentence	Words providing context
	Position of sentence	The first or second sentence
	Processing	Inserting words after the sentence to provide information
	Inserting words	"What is that?"†1 or "Give me more details"†2
Rule 2:	*Repetition questions*	
	Objective	To promote listener's understanding by repeating questions asked
	End of sentence	Question
	Position of sentence	The last sentence
	Processing	Inserting words after the sentence represents asking a question
	Inserting words	"What do you think about that?"†3
Rule 3:	*Simple response*	
	Objective	Allowing time for the listener to understand by inserting a simple response
	End of sentence	No use
	Position of sentence	The second sentence
	Processing	Inserting simple responsive words after the second sentence
	Inserting words	"Yes"or "Uh-huh"†4

†1 "Sore ha nandesuka?"in Japanese.
†2 "Motto kuwashiku osiete kudasai"in Japanese.
†3 "Dou omoimasuka?"in Japanese.
†4 "Hai"or "Un"in Japanese.

When an opinion has been submitted with a still picture, the picture is displayed in the center of the screen. Figure 2 shows the POC caster display. On the screen, two agents with face photos and animated human bodies introduce each community member's opinions in turn. The two agents provided information as follows: (1) Agent 1 reads out the title, and this allows us to get a summary of the topic. (2) Agent 2 introduces the text of a previous opinion sentence (The original text form is divided into two parts by a period). (3) Agent 1 comments on the previous text based upon certain rules. (4) Agent 2 proceeds to the later parts of the text.

2.1 Transformation Rules of POC caster

The original text is transformed into a conversational representation by applying transformation rules constructed based on the end of each opinion sentence and the position of each sentence in the text. The end of a sentence can indicate two types of expression: (A) declarative expressions to indicate information related to interests, clarify a situation, or indicate hearsay; and (B) interrogative expressions. In Japanese, we can grasp the meaning of a passage by focusing our attention on the statement's end because that is often where the verb is located (unlike English). Sentences stating opinion could be "initial", "middle", or "final"ones. For example, if there are only two sentences, no middle sentence exists, and if there is only one sentence, it is treated as the final one. Using these cues, we have implemented three transformation rules for conversational representation (Table 1). Rule 1 shows that the following sentence gives the listener a detailed description by presenting contextual information. Rule 2 makes the listener pay attention to the topic by repeating questions from the previous sentence. Rule 3 allows the listener to gain time to understand the meaning by inserting a simple response (for more details on the technical aspects of POC caster, see Kubota et al., 2002; Kubota, Yamashita, & Nishida, 2002).

3 Experiment 1

In this experiment, we evaluated the transformation rules of POC caster through a psychological factorial experiment. The main purpose of this experiment was to examine whether the con-

Table 2: Examples of Transformation to Conversational Representation

Original text
The time when genes can be patented has come. Scientists able to patent a disease-causing gene can prohibit other scientists from studying it. Medical development will be obstructed by such patents being applied.
Conversational representation
A: The time when genes can be patented has come.
B: What is that? †
A: Scientists able to patent a disease-causing gene can prohibit other scientists from studying it. Medical development will be obstructed by such patents being applied.

†Words inserted using Rule 1.

versational representation generated by POC caster, compared to text representation, could aid sentence comprehension in relation to sentence length. We hypothesized that conversational representation would enhance comprehension of the presented sentence. More specifically, when the presented sentence was short, we expected the effect of conversational representation to be lower than when the sentence was long. Moreover, we explored the effect of words inserted to generate the conversational representation.

3.1 Method

3.1.1 Design and Participants

A 2 x 2 within-subjects design was used with one factor of representational form (single speech and conversation) and a second factor of sentence length (long and short). Twenty-four people (21 males and 3 females) between 21 and 41 years of age (M = 24.92) participated, and were paid for their cooperation.

3.1.2 Materials

Fifty long sentences (160-200 characters) and fifty short sentences (50-100 characters) were selected from three-hundred sentences taken from opinions submitted to POC, newspaper articles, or dictionaries. Thirty long sentences and thirty short sentences were pooled as a stimulus set. Sentences were excluded from this set when there was a significant difference between the time needed for participants to read the sentence from text and the time needed for the speech synthesis system to read out it aloud. We applied the POC caster transformation rules to sixty sentences and generated conversational sentences such as inserted words that aided understanding of the context (e.g., "What is that?") and simple responses to allow time for understanding (e.g., "Yes.") (see Table 2 for an example).

3.1.3 Procedure

Participants were tested individually in a single session lasting about 60 minutes. Each session included instruction, a practice trial, experimental trials, and completion of a questionnaire. The participants were seated in front of a personal computer (Pentium III, 700 MHz, Windows 98), and were given written instructions specifying the purpose of this study (i.e., to evaluate an agent technology series). During each trial, the stimulus sentence was provided through headphones, and the participant judged the sentence comprehensibility on a seven-point scale ("1: easy"to "7: difficult") by responding with a mouse click as quickly as possible without making errors. Participants were instructed not to allow the speech synthesis quality to affect their judgment.

A total of 32 sentences (16 long and 16 short) were selected at random from a stimulus set for each participant. The trials consisted of two blocks of 16 sentences (8 long and 8 short) with a two-minute break between blocks. The sentence order was random in each block for each

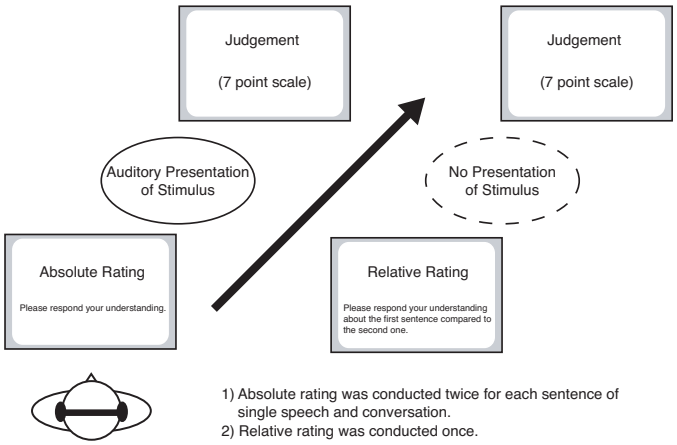


Figure 3: Schematic of a trial sequence in Experiment 1.

participant. Participants were asked to judge comprehensibility three times per trial (Figure 3). In the absolute rating, they judged each of two sentences. (The sentence order was random for each participant.) In the relative rating, participants compared two sentences and judged whether the first or second sentence was easier to understand.[1] The conversational representation included two kinds of sentence. One included words intended to help the participant understand the context (rich context), the other consisted of simple responsive words used to obtain time for understanding (poor context). Each sentence type appeared an equal percentage of the time.

3.2 Results and Discussion

For all of the analyses reported in this paper, the statistical rejection level for significance was set at $p < .05$ unless otherwise indicated. Data from four participants were excluded from all following analyses - one because of a computer problem, and three because of their failure to comply with the instructions. In consequence, we analyzed data from 23 participants for the absolute rating and 20 for the relative rating. Response time (RT) was measured through the participants' mouse clicking, but it is not reported here because we are not fully confident of its reliability. (The RT results we obtained, however, were consistent with the patterns of the rating score results.)

Absolute rating. Average rating scores are summarized in Table 3. Results from an analysis of variance (ANOVA) revealed a significant main effect of sentence length, $F(1, 22) = 34.36$, and a significant interaction between the sentence length and the representational form, $F(1, 22) = 19.32$. A posteriori Tukey tests showed that the short sentences were easier to understand than long sentences both in single speech and conversational representation. Moreover, the tests indicated that for a long sentence the conversational representation was easier to understand than single speech, whereas there was no significant difference for a short sentence. Table 4 shows the effect of inserted words (rich vs. poor). ANOVA results showed a significant main effect of sentence length, $F(1, 22) = 15.09$, and that of inserted words, $F(1, 22) = 43.45$. That is, the enhancement of comprehension was greater when the sentence was short ($M = 5.27$) than when it was long ($M = 4.79$), and it was greater when inserted words were of rich context ($M = 5.37$) rather than poor context ($M = 4.69$).

Relative rating. Table 5 shows average scores of the relative rating for comprehension. We conducted a three-way ANOVA with sentence length (long vs. short), inserted words (rich vs. poor), and judgment sentence (single speech-conversation vs. conversation-single speech) as within-subject variables. The ANOVA revealed a significant main effect of sentence length ($F(1, 19)$

[1] When the single speech sentence was judged, the stimulus presentation was done in the order of the single speech sentence and then the conversational sentence, and participants rated the comprehensibility for the single speech sentence compared to the conversational one.

Table 3: Average Rating Scores for Representational Form in Absolute Rating as a Function of Sentence Length.

	Sentence length	
	Long	Short
Single speech	4.82	4.55
	(.64)	(.76)
Conversation	5.25	5.39
	(.67)	(.81)

Note. $N = 23$. Standard deviation (SD) values are given in parentheses.

Table 4: Average Rating Scores for Inserted Words as a Function of Sentence Length.

	Sentence length	
	Long	Short
Inserted words		
Poor	4.47	4.91
	(.77)	(.87)
Rich	5.11	5.63
	(.75)	(.63)

Note. $N = 23$. Standard deviation (SD) values are given in parentheses.

= 4.58), a significant interaction between sentence length and judgment sentence ($F(1, 19) = 13.82$), and a significant interaction between inserted words and judgment sentence ($F(1, 19) = 10.10$). Subsequent analysis of the interaction between the sentence length and judgment sentence revealed that when participants were tested on single speech-conversation, they rated more highly with short sentences ($M = 4.38$) than long ones ($M = 4.54$), $F(1, 19) = 15.93$, and when tested on conversation-single speech, they scored more highly with long sentences ($M = 4.54$) than short ones ($M = 4.26$), $F(1, 19) = 5.10$. Also, while there was no significant difference in the judgment sentence when the stimulus was a short sentence, when the stimulus was a long sentence, the rating score was higher in conversation-single speech ($M = 4.54$) than single speech-conversation ($M = 3.79$), $F(1, 19) = 9.15$. Subsequent analysis of the interaction between inserted words and judgment sentences showed that there was no significant difference between single speech-conversation ($M = 4.17$) and conversation-single speech ($M = 4.20$) when the inserted words were of poor context; however, when the inserted words were of rich context, rating scores for conversation-single speech ($M = 4.60$) were higher than those for single speech-conversation ($M = 4.00$), $F(1, 19) = 8.60$. In addition, the effect of the inserted words was not significant in single speech-conversation, but was in conversation-single speech, and participants scored higher when sentences with rich context ($M = 4.60$) were presented than when sentences with poor context were presented ($M = 4.20$), $F(1, 19) = 6.86$.

These results may be summarized as follows. The conversational representation generated by the POC caster transformation rules promotes comprehension more effectively when the sentence length is long than when it is short; inserting words having rich context information has a stronger effect on our comprehension than simple responses. Although it is important that we have demonstrated when conversational representation will be beneficial, note that the reliability of our results is limited for three reasons. First, the voice provided by the speech synthesis system was artificial with respect to accent, intonation, rhythm, and so on. We cannot exclude the possibility that the unnatural voice affected the participants' judgment even though we instructed them not to pay attention to the speech synthesis quality. Second, the pauses between sentences were picked at

Table 5: Average Relative Rating Scores for Sentence Judgment as a Function of Sentence Length and Inserted Words.

	Sentence length			
	Long		Short	
Inserted words	Poor	Rich	Poor	Rich
Single speech	3.78	3.80	4.56	4.21
- Conversation	(.75)	(.59)	(.93)	(.77)
Conversation	4.30	4.79	4.10	4.41
- Single speech	(.88)	(.73)	(.63)	(.71)

Note. $N = 20$. Standard deviation (SD) values are given in parentheses.

random by the speech synthesis system and were not manually controlled. Third, there may have been methodological problems in our procedure. For example, participants might have found it difficult to separately judge the absolute and relative ratings. These problems were resolved in Experiment 2, which is described below.

4 Experiment 2

In Experiment 1, we examined the validity of the POC caster transformation rules. Our results indicated that conversational representation was superior when the sentence length was long and when inserted words had rich context information. A participant's understanding, though, is likely to depend on that person's knowledge level regarding the topic. After the trials in Experiment 1, some participants made comments such as "I could easily understand the sentence when I knew the topic very well."The effect of context should be maximized when we know the topic well. Our main purpose in Experiment 2 was to investigate the relationship between a participant's knowledge level and conversational representation by using a forced-choice task. The idea was that if a participant knew little about the topic, conversational representation should aid understanding. On the other hand, if a participant was knowledgeable about the topic, conversational representation should not strongly affect understanding.

4.1 Method

4.1.1 Design and Participants

We used a 2 x 2 x 2 mixed design with one factor of representational form (single speech vs. conversation) as a within-subject variable, a second factor of knowledge level (high vs. low) as a within-subject variable, and a third factor of modality (visual vs. auditory) as a between-subject variable. Seventy-eight people (45 female and 33 male) between 19 and 35 years of age ($M = 22.21$, SD = 3.55) participated in this experiment and were paid for their cooperation. Participants were randomly assigned in equal numbers to two groups (one for the visual and the other for the auditory condition).

4.1.2 Materials

A total of 108 sentences were selected from newspaper and magazine articles within the range of 160-200 words (corresponding to the long-sentence length of Experiment 1). The sentences included 36 topics (e.g., environment-related issues, fortune-telling, horse races, football, proverbial phrases, insects). We applied the POC caster transformation rules and generated conversational sentences with inserted words to aid understanding of the context (e.g., "What is that?") and simple response words to gain time to understand (e.g., "Yes."). The voice stimuli for the auditory condition were recorded with a human male voice (Standard Japanese). The conversation stimuli consisted of male and female voices. The pauses between sentences were 1500 ms long, and pauses

following a comma were 700 ms long. The speech was at a natural speed. These modifications were made to resolve the above-mentioned problems in Experiment 1.

4.1.3 Procedure

Participants were tested individually in a single session lasting about 90 minutes. The procedure consisted of three parts: a knowledge task, experimental trials, and questionnaire completion.

In the knowledge task, participants were asked to rate their level of familiarity and interest regarding words presented on a personal computer (Pentium III, 700 MHz, Windows 98) on a five-point scale ("1: not at all"to "5: very high"). Thirty-six words were presented in the center of a computer screen one at a time and in a random order for each participant. From the results of the knowledge task we then selected 12 words (6 words representing well known topics and 6 words representing unfamiliar topics) as theme words for each participant.

Next, in the experimental trials, participants were instructed to judge which of two sentences presented by a personal computer was easiest to understand. They were also told to press a particular key on the keyboard as quickly as possible when the word "Judgment"appeared on the screen (Figure 4). They were asked to press the "1"key if the first sentence was easiest to understand, and to press the "|"key if the second sentence was easiest. We used a forced-choice method. Stimuli under the visual condition were presented in the center of the computer screen, whereas those under the auditory condition were presented though headphones. Under the auditory condition, turn-switching in conversations could be clearly recognized because of differences in the speakers' voices. Under the visual condition, "A: "and "B: "were displayed before the sentences to visually indicate the conversational form. Two sentences were presented as one pair for each trial. The trials consisted of 2 blocks of 18 trials, with a three-minute break between blocks. Each sentence pair was presented in a random order within a block. The total presentation time required to display characters or recall the recorded voices was adjusted so that the presentation time was not markedly different between the two conditions. There were two practice trials using sentences not used in the experimental trials.

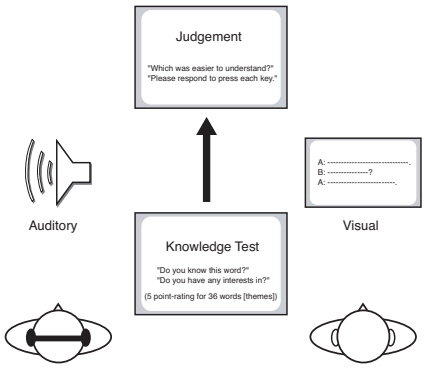


Figure 4: Schematic of a trial sequence in Experiment 2.

After the experimental trials, participants were asked to answer the following questions in an open-ended style: "What kind of reason led you to select the sentence?", and "What kind of impression did you have?"

4.2 Results and Discussion

Data from 36 participants were excluded from all following analyses because of the knowledge task criteria. That left 21 participants (7 male and 14 female) for the visual condition (age: M = 23.57, SD = 4.39) and 21 participants (10 male and 11 female) for the auditory condition (age:

Table 6: Average Probability of Selecting the Representational Form as a Function of Knowledge Level and Modality in Experiment 2.

	Knowledge level	
	High	Low
	Visual presentation	
Single speech	.67	.57
	(.22)	(.17)
Conversation	.33	.43
	(.22)	(.17)
	Auditory presentation	
Single speech	.58	.46
	(.22)	(.22)
Conversation	.42	.54
	(.22)	(.22)

Note. $N = 21$ for both the visual condition and the auditory condition. The standard deviation (SD) is given in parentheses.

$M = 21.76$, $SD = 3.06$). Our dependent measures were the number of times that participants selected a particular sentence and the response latency.

The probability of selection is shown in Table 6. A 2 (representational form) x 2 (knowledge level) x 2 (modality) mixed ANOVA for data with inverse sine transformation revealed a significant interaction between the representational form and the knowledge level, $F(1, 40) = 16.36$ and a marginally significant interaction between the representational form and modality, $F(1, 40) = 2.99$, $p < .10$. The former interaction only is reported here. Post hoc analysis of the former interaction revealed no reliable difference in the selection of sentences when the level of knowledge was low; in contrast, for sentences where the level of knowledge was high, selection was clearly higher for simple speech ($M = .54$) than for conversation ($M = .35$), $F(1, 40) = 13.17$. Also, when we tested a simple main effect of the representational form, we found that sentences for which the participant's level of knowledge was high were selected significantly more often than those for which the level of knowledge was low under the simple speech condition, $F(1, 40) = 15.71$; under the conversation condition, a low level of knowledge led to higher selection than did a high level of knowledge, $F(1, 40) = 16.52$.

Table 7 shows the mean reaction time (RT) of the selective responses. A 2 (representational form) x 2 (knowledge level) x 2 (modality) mixed ANOVA for data with logarithmic transformation revealed that there was a significant second-order interaction, $F(1, 40) = 6.18$. Thus, we computed a 2 (representational form) x 2 (knowledge level) ANOVA for each modality condition since we were interested mainly in the relationship between the knowledge level and representational form. The interaction between the representational form and knowledge level was reliable for the auditory condition, $F(1, 20) = 8.58$, but there was no significant effect for the visual condition. As to the significant interaction for the auditory condition, participants with a high level of knowledge were faster to respond to a sentence of simple speech than to one of conversation, $F(1, 20) = 3.47$, $p < .10$. On the other hand, participants with a low level of knowledge were faster to respond to conversation than to simple speech, $F(1, 20) = 4.09$, $p < .10$. Under the simple speech condition, there was no significant difference that depended on the participant's level of knowledge; under the conversation condition, though, participants with a low level of knowledge responded to sentences more quickly than those with a high level of knowledge, $F(1, 20) = 9.23$.

In Experiment 2, we examined the relationship between conversational representation and knowledge level. Our results are as follows. (1) The conversational representation had a beneficial effect when participants were knowledgeable about the topic; however, conversational representation had no effect when participants had little knowledge of the topic. (2) Under the conversational condition, participants with no knowledge of the topic were more likely to consider the sentence

Table 7: Average Response Time when Selecting the Representational Form as a Function of Knowledge Level and Modality in Experiment 2

	Knowledge level	
	High	Low
	Visual presentation	
Single speech	973.63	1058.04
	(483.37)	(579.19)
Conversation	1057.68	1090.53
	(540.49)	(517.33)
	Auditory presentation	
Single speech	752.05	1025.03
	(316.79)	(675.65)
Conversation	1063.97	800.89
	(812.73)	(433.49)

Note. $N = 21$ for both the visual condition and the auditory condition. Response times are given in milliseconds. The standard deviation is given in parentheses.

as easier to understand than the knowledgeable participants. (3) The response time was shorter with conversational representation when participants had little knowledge about the topic. These results indicate that the effect of conversational representation depends on the user's relevant knowledge.

5 General Discussion

What kinds of information agents should give, when they should speak, and how they should speak are important issues. In this study, we have investigated what representational form is needed to promote user comprehension in two experiments with a factorial design. We have shown the beneficial effect of conversational representation in a long sentence (in Experiment 1) and for a person lacking relevant knowledge (in Experiment 2) when software agents provide information in a broadcasting system. These results support our basic hypothesis that conversational representation aids our comprehension. Related work includes a series of studies done by Mayer and his colleagues (Mayer, 2003; Moreno, Mayer, Spires, & Lester, 2001; Moreno & Mayer, 2002) in which the type of presentation was examined. For example, Moreno and Mayer (2001) experimentally varied whether the agent's words were presented as speech or as on-screen text. Their results showed that people learned better when words were presented as speech rather than as on-screen text. The main difference between our study and their work is in the conversational representation. In that sense, we are the first to demonstrate that there are specific conditions under which conversational representation, compared to text representation, aids comprehension.

We found in Experiment 2 that conversational representation allowed a person without relevant knowledge to grasp the sentence focus and the connection between sentences. This was because users could easily process the sentence since the text was divided into small parts. In the field of cognitive psychology, Kintch (1998) has proposed a "macro-structure model", which is a semantic structure of the entire sentence, and a "situation model", which is connected to knowledge. He has also suggested a global model for text comprehension, which consists of bottom-up and top-down processing (see also van Dijk & Kintch, 1983). According to Kintch and his colleagues, this macro-structure model means that text comprehension depends on making an exact representation (text-base) for the content of a text, from analysis of each word in a text to integrating higher meanings, and integrating this into the higher meaning of the whole text (macro-structure). This can be considered "learning of text"processing. On the other hand, the situation model is aimed at understanding the meaning of an object that is explained in a text, relating the content to our knowledge. This can be considered "learning from text"processing. These forms of processing are

not thought of as alternatives, but as forms that run in parallel.

When we know a topic well, our processing can focus on checking for current information relative to our existing knowledge because the situation model can be readily formulated. On the other hand, when we lack relevant knowledge, our processing is dominated by the bottom-up processing needed for superficial understanding because the situational model cannot be formulated or will be inaccurate. Applying these models to the conversational representation relationship allows a person without knowledge regarding a particular topic to grasp the issue and connection between sentences. Inserting conversational representation in this case supports text-based processing, and facilitates text understanding. On the other hand, conversational representation has no effect on understanding for a person who is knowledgeable because that person is able to create the situation model without help; that is, such users only have to verify the current information relative to their existing knowledge. Thus, conversational representation supports text-base processing only for those who cannot formulate a sufficient situational model.

It can also be said that conversational representation is a kind of "advance organizer" (Ausubel, 1963; 1978) in the sense that specific contextual information is added in advance. Ausubel (1963) defines an advance organizer as the "anchoring foci for the reception of new material. It allows the learner to recall and transfer prior knowledge to the new information being presented. This theory is based on the idea that learning is facilitated if the learner can find meaning in the new information. If a connection can be made between new information and previous knowledge, the learning experience will become more meaningful to the learner. Studies have shown that comprehension and memory of a text is promoted by presenting particular information such as a title (Bransford & Johnson, 1972), a summary (Bromage & Mayer, 1986), a perspective (Pitchart & Anderson, 1977), or an illustration (Waddili & McDaniel, 1988) before the text. Do any qualitative differences added to an original text affect our comprehension? If so, what kind of difference is needed? Further research is needed to clarify this point by manipulating the factor of inserted words more precisely.

The success of text comprehension or learning is likely to depend on the interaction between the characteristics of the learner and the properties of the text itself (e.g., amount of information provided about the topic, connection between sentences, clarity of wording). Although we have studied the effect of one learner characteristic (knowledge level) and one text property (sentence length) in this experiment, many factors remain to be examined. Further research is needed to investigate these factors and the relationships between them.

References

Azechi, S., Fujihara, N., Sumi, K., Hirata, T., Yano, H., & Nishida, T. (2000). Public Opinion Channel: A challenge for interactive community broadcasting. In Ishida, T. (eds.), *Digital Cities: Experiences, Technologies and Future Perspectives, Lecture Notes in Computer Science*, pp. 427-441. Springer-Verlag.

Ausubel, D. (1963). *The Psychology of Meaningful Verbal Learning.* New York: Grune & Stratton.

Ausubel, D. (1978). In defense of advance organizers: A reply to the critics. *Review of Educational Research, 48,* 251-257.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning & Verbal Behavior, 11,* 717-726.

Bromage, B. K., & Mayer, R. E. (1986). Quantitative and qualitative effects of repetition on learning from technical text. *Journal of Educational Psychology, 78,* 271-278.

Cassel, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., & Yan, H. (2000). Human conversation as a system framework: Design embodied conversational agents. In Cassell, J. et al. (eds.), *Embodied Conversational Agents*, pp. 29-63. Cambridge, MA: MIT Press.

Fukuhara, T., Chikama, M., & Nishida, T. (2003). Supporting an Experiment of a Community Support System: Community Analysis and Maintenance Functions in the Public Opinion Channel. In M. Huysman, E. Wenger, and V. Wulf (eds.), *Communities and Technologies*, pp. 347-367. Dordrecht: Kluwer Academic Publishers.

Kintch, W. (1998). *Comprehension: A Paradigm for Cognition.* Cambridge University Press.

van Dijk, T. A., & Kintch, W. (1983). *Strategies for Discourse Comprehension.* New York: Academic Press.

Kubota, H., Yamashita, K., Fukuhara, T., & Nishida, T. (2002). POC caster: Broadcasting agent using conversational representation for Internet community. *Transactions of the Japanese Society for Artificial Intelligence, 17,* 313-321. (in Japanese with English abstract)

Kubota, H., Yamashita, K., & Nishida, T. (2002). Conversational contents making comment automatically. In E. Damiani, R. J. Howlett, L. C. Jain, & N. Ichalkaranje (Eds.), *Frontiers in Artificial Intelligence and Applications: Vol. 4. Knowledge-based Intelligent Information Engineering Systems & Allied Technologies, Part 2*, pp. 1326-1330. Amsterdam: IOS Press

Mayer, R. E., Dow, G. T., & Mayer, S. Multimedia Learning in an Interactive Self-Explaining Environment: What Works in the Design of Agent-Based Microworlds? *Journal of Educational Psychology, 95,* 806-812.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19,* 177-213.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598-610.

Pitchart, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology, 69,* 309-315.

Rickel, J., & Johnson, W. L. (1998). STEVE: A pedagogical agent for virtual reality. In Syncara, K. P. and Wooldridge, M. (Eds.), *Proceedings of the 2nd International Conference on Autonomous Agents (AGENTS 98)*, pp. 332-333, New York, May 9-13 1998. ACM Press.

Waddili, P. J., & McDaniel, M. A. (1988). Illustrations as adjuncts to prose: a text-appropriate processing approach. *Journal of Educational Psychology , 80,* 457-464.