

Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002

Saeko Nomura* Toru Ishida[†] Mika Yasuoka*

*Department of Social Informatics, Kyoto University

[†]Japan Science and Technology Corporation

{saeko, mika}@kuis.kyoto-u.ac.jp, ishida@i.kyoto-u.ac.jp

INTRODUCTION

Toward the dramatic Internet diffusion in Asia-Pacific countries [1], the intercultural collaboration support system, which surpasses the sense of values and language differences, is necessary to be constructed. Differing from the face-to-face communications, to the computer-mediated communications, machine translation services can be easily applied. This technological advancement drastically increases availability of intercultural collaboration beyond the language barrier.

Although natural language processing researchers have conducted rigorous research on machine translation for years ([2][3][4] etc.), the translation quality is hardly adequate to be applied to practical worksites. The preceding studies evaluate machine translation on written documents, and do not take into account “interaction factor” to refine translation quality. We, on the other hand, apply machine translation to human-to-human collaboration, and try to analyze the interactive translation refinement procedures implemented between humans and between human and machines.

INTERCULTURAL COLLABORATION EXPERIMENT 2002

Intercultural Collaboration Experiment 2002 (ICE2002) is an experimental project among Asian universities. In this project, 31 students from Kyoto University (Japan), Shanghai Jiaotong University (China), Seoul National University and Handong University (South Korea), and University of Malaya (Malaysia) develop software over the Internet. Since this experiment pursuit collaboration among heterogeneous groups across country borders, participants never see each other and communicate in their mother languages supported by machine translation.

In ICE2002, multilingual communication tools, named TransWEB and TransBBS were provided to participants (see Fig. 1). These tools incorporate translation services¹ among Chinese, Japanese, Korean, Malay, and English. TransBBS, a multilingual bulletin board system, is utilized as a daily discussion space. TransWEB enabled participants to browse software development documents in their mother languages. Discussions in this experiment are held only in TransBBS and TransWEB.

The first track of ICE2002 was conducted from May to July

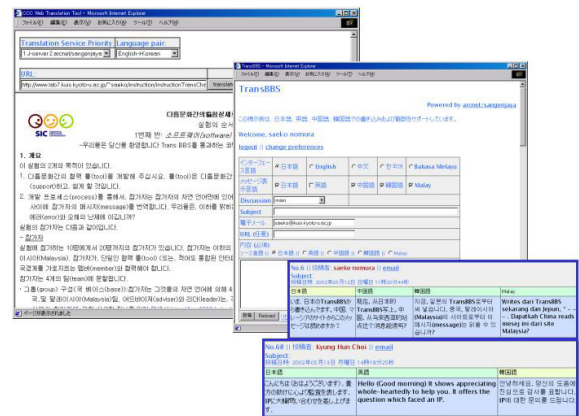


Figure 1. Communication Space in ICE2002:
TransWEB and TransBBS

2002. To ensure smooth communication and software development, the experimental schedule was divided into the following two phases (4 weeks respectively):

- *Software Design Phase*: Intercultural collaboration software is designed. The goal of this phase is to submit a system design proposal to implement software.
- *Software Implementation Phase*: Software based on the design proposal is implemented. The goal of this phase is to complete and release an intercultural collaboration tool.

LESSONS LEARNED

In ICE2002, intercultural collaboration software, which consists of Web-based email, SMS, and search engine, is submitted as an outcome by a multinational team. All modules of this software incorporate translation services.

To analyze translation-mediated collaboration on TransBBS, the conversation [5] and content analysis [6] were applied. We found different communication patterns between phases:

- *Software Design Phase*: Postings about translation qualities, and interactions for avoiding misunderstandings among participants.
- *Software Implementation Phase*: Active Q&A relevant to the software development, regardless of translation qualities.

¹ Translation services were provided by arcnet/sanganjaya (<http://sanganjaya.arc.net.my/index-e.html>) and J-server (<http://www.j-server.com/index.shtml>) via Internet.

In the *Software Design Phase*, participants tried to adapt themselves to the machine translation service with the intention of correctly conveying their remarks to the participants from other countries. The following three methods were utilized to correct the translation errors in most cases: 1) Creating a new translation test discussion thread and repeating the practice postings, 2) Browsing posted messages in five languages to check if posted messages are translated properly, 3) Asking other participants if a posted message is comprehensible in other languages.

With respect to the third method, for instance, we observed following posting on TransBBS.

"We want to know whether you could understand the my contribution japana-148 (message ID) or not. We, please contribute using your language the part that you understood. I want you to also contribute us to japana-148's opinion." by a participant from Japan Group (written in Japanese and translated by arcnet/sangenjaya)

This message was posted by a Japan Group member. The noticeable phenomenon here is a particular Japanese phrase. In the process of TransBBS usage, the contributor of this message understood that the natural Japanese sentences (such as omitting subjects) mislead the appropriate translation results. He started adapting to the machine translation capability and using specific Japanese phrases (such as adding subjects). In the *Software Design Phase*, we observed some interactive translation refinement procedures implemented between humans and between human and machines. These procedures helped participants to construct social capital among them.

In the *Software Implementation Phase*, some participants began to conduct active technical and practical discussions in the context of software implementation regardless of the translation quality (for example, see Fig.2). These active participants confirmed their collaboration results (mutual understandings and consensus) in the submitted program source code on CVS.

CONCLUSION

In this paper, we observed the communication patterns on ICE2002, and considered the possibility of machine-translated collaborative works beyond cultural borders. By observing the first track of ICE2002, we found that machine translation services indeed supported intercultural collaboration in the software development process. The approach of this research is to use translation function in collaborative context. We continue to analyze the human-to-machine interaction, and pursue the possibility of 'translation-mediated groupware' which can be applied to the multinational project team activities.

M:	"TransMail is a good idea. But how will the content be translated and if the mail has an attachment..will that be translated too?"
J:	"The appending file will not be translated. Because it is very difficult."
K:	"You can be seen or there is not a problem of the sticking file in the case of TransMail importantly like this (so). It sees that will be possible sufficiently if passing through server to raise by playing Wrapper role to pass through SMTP."

<p>" I could understand your message. There are some kinds in the TransMail's mounting.</p> <p>Type 1: Sender --> ServerA --> ServerB --> Receiver SMTP SMTP SMTP</p> <p>TransMail is on ServerA and works as WRAPPER of SMTP.</p> <p>Type 2: J: Sender --> ServerA --> ServerB --> Receiver HTTP SMTP SMTP</p> <p>TransMail works as WebApplication on ServerA. TransMail translates a message and gives SMTPD that message.</p> <p>Type 3: Sender --> ServerA --> ServerB --> Receiver HTTP RSYNC HTTP</p> <p>TransMail works as WebApplication on ServerA and SearverB. This is the same as the mounting of the present TransBBS. We decided to adopt Type 3 in this time."</p>	
--	--

Figure 2. Intercultural Collaboration in the *Software Implementation Phase*

- M: Malaysia Group Member's posting (written in English)
- J: Japan Group Member's posting (written in Japanese)
- K: Korea Group Member's posting (written in Korean)

REFERENCES

- [1] M. Pastore, "Asian Internet Market to Surpass United States," CyberAtlas, August 7, 2001, http://cyberatlas.internet.com/big_picture/geographics/print/0,,5911_861481,00.html
- [2] W. J., Hutchins and H. L. Somers, *An Introduction to Machine Translation*, Academic Press, 1992.
- [3] D. Jurafsky and J. H. Martin, *Speech and Language Processing - An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Prentice Hall, 2000.
- [4] B. Maegaard Ed., *Proceedings of MT Summit VIII*, 2001.
- [5] P. T. Have, *Doing Conversation Analysis: A Practical Guide (Introducing Qualitative Methods)*, Corwin Press, 1999.
- [6] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, Sage Publications, 1981.