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

Saeko Nomura[†] Toru Ishida[†‡]*
Mika Yasuoka[†]

Naomi Yamashita[*]
Kaname Funakoshi[*]

Dept. of Social Informatics, Kyoto University[†]
Japan Science and Technology Corporation[‡]
{saeko, mika}@kuis.kyoto-u.ac.jp
ishida@i.kyoto-u.ac.jp

NTT Communication Science
Laboratories, NTT Corp[*]
{naomi, kf}@cslab.kecl.ntt.co.jp

Abstract

This paper reports about the Intercultural Collaboration Experiment (ICE2002), and discuss the effectiveness of machine translation for task oriented multinational teams. In ICE2002, students from China, Japan, Korea, and Malaysia developed software in collaboration over the net. In the experiment, participants communicated in their mother languages using machine translations. By analyzing communication log data, we found that, in the first place, participants repeated repairs to overcome translation errors to understand each other. Gradually, they started to hold active discussions regardless of translation qualities. We confirmed that machine translation has enormous potentialities to break language barriers in the multinational collaboration process.

1   Introduction

Toward the dramatic Internet diffusion in Asia-Pacific countries (Pastore, 2001), 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 (Maegaard Ed., 2001, 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.

[†] Honmachi Yoshida, Sakyo-ku, Kyoto, Japan 606-8501
[‡] 366 Ichino Funairi-cho, Kawaramachi Nijo sagaru,Nakagyo-ku, Kyoto, Japan 604-0924
[*] 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan 619-0237

This paper reports about the result of Intercultural Collaboration Experiment 2002 to consider the applicability of machine translation to multinational collaborative works.

2   Intercultural Collaboration Experiment 2002

Intercultural Collaboration Experiment 2002 (ICE2002)[1] is an experimental project among Asian universities. In this project, 32 students from Kyoto University (Japan), Shanghai Jiaotong University (China), Seoul National University and Handong University (South Korea), and University of Malaya (Malaysia) developed software over the Internet (see Figure 1). Since ICE2002 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 TransBBS and TransWEB were provided to participants (see Figure 2). These tools incorporate translation services [2] 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. Communications in ICE2002 are held only on TransBBS and TransWEB.
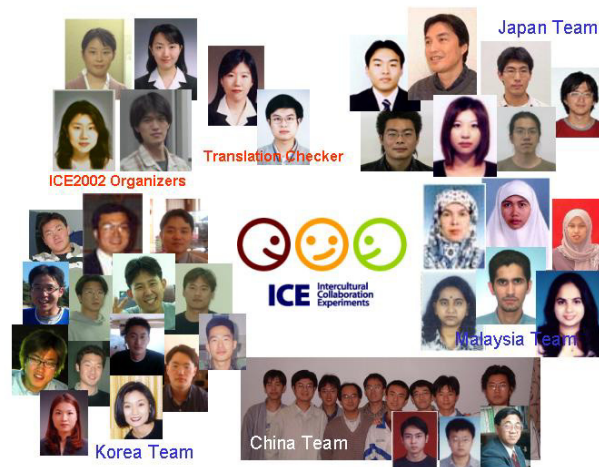

Figure 1: ICE2002 Participants

ICE2002 had two tracks. The first track was conducted from May to July, and the second one from October to December. To ensure smooth communication and software development, each track was divided into the following two 4-week-phases:
- 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.

3   Communication Patterns Using Machine Translation in ICE2002

In ICE2002, intercultural collaboration software, which consists of Web-based email, SMS, search engine, is submitted as an outcome by a multinational team[3]. In this section, we observe how participants achieved their goal with communications in a noisy media.

---

[1] http://ice.kuis.kyoto-u.ac.jp/ice2002/

[2] Translation services were provided by arcnet/sangenjaya (http://sangenjaya.arc.net.my/index-e.html) and J-server (http://www.j-server.com/index.shtml) via Internet.

[3] Chat module which was implemented by Korean students could not be integrated to the outcome.
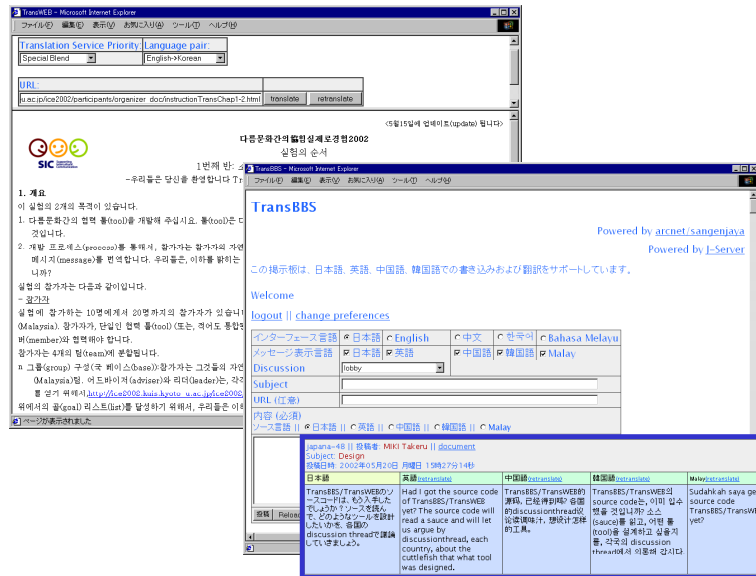
Figure 2: Multilingual Communication Tools Used in ICE2002:
TransWEB (left) and TransBBS (right)

To analyze translation-mediated collaboration on TransBBS, the conversation (Have, 1999) and content analysis (Krippendorf, 1981) were applied. We found different communication patterns between phases:

- Software Design Phase: Translation repair trials before posting, postings about translation qualities, and interactions for avoiding misunderstandings among participants were observed.
- Software Implementation Phase: Active questions and answers about software development, regardless of translation qualities, were confirmed.

In the **Software Design Phase**, participants concerned about translation quality and tried to explore ways to convey their remarks correctly to others from different countries. The following two methods were observed:

- *Other-initiated repair (OIR)*: Collaborative translation-error repair process between message contributors and receivers to achieve sharing common knowledge. The repair process is initiated by a message receiver's reaction.
- *Self-initiated repair (SIR)*: Before posting a message, a message contributor repeats repairing by him/herself to refine translated results.

In ICE2002, we confirmed *repair organization* (Schegloff, 1992) especially in the process of translation-error eliminations. With respect to *other-initiated repair* (*OIR*), for instance, we observed following interaction on TransBBS.

Iizawa (Japanese) :	"I write for the first time. It is the Iizawa of its Japanese team. *It is asked properly*."
HaoLei (Chinese) :	"I estimate the friend of Japan said exactly "*Please concern more*" meaning?"
Miki (Japanese) :	"It is as it says. It read Japanese translation of the meaning which HaoLei supposed. It agree with Japanese which Iizawa wrote."

* Messages are written in Japanese and Chinese respectively and translated into English by arcnet/sangenjaya.

Here, HaoLei (Chinese) pointed out an estimated translation error and guessed a proper message. By the answer from Miki (Japanese), we can find that they shared a proper common knowledge. This kind of *OIR* was often seen in the beginning of the experiment.

TransBBS has a function which enables message contributors to check translation results and to refine their remarks before posting. With this function, most participants conducted *self-initiated repair* (*SIR*). Table 1 shows number of posts and *SIRs* of top five active contributors on TransBBS. Tomoko, a Japanese participant, tried to conduct 7.5 times of *SIR* per one post on average.

Table 1: Number of *SIR* on TransBBS: Software Design Phase

Name	# post / *SIR*	Average #*SIR* / one post
Saeko (Japanese)	69 / 234	3.4
Bikesh (Malay)	66 / 112	1.7
Tomoko (Japanese)	**51 / 395**	**7.5**
Yoshiyuki (Japanese)	35 / 170	4.9
LimSanny (Korean)	28 / 71	2.5

Following example is *SIR* by Tomoko. After she tried *SIR* five times, she posted a message on TransBBS. The message posted after *SIR* is easily understandable compared to the initial one.

Initial message before repairing : "My sentence **is also and it is not translated well**. **Will it begin** from the study of Japanese?"

After five times self-initiated repair…

Posted message on TransBBS : "My sentence **is not translated effectively either. I need to begin** the study of Japanese once more."

* Messages are written in Japanese translated into English by arcnet/sangenjaya.

In the process of *SIR*, Tomoko understood that the natural Japanese sentences (such as omitting subjects) mislead appropriate translation results. Then she started adapting to the machine translation capability and using specific Japanese phrases (such as adding subjects).

In this way, in the Software Design Phase, we observed some interactive translation refinement

Puja (Malay):	"TransMail is a good idea. But how will the content be translated and if the mail has an attachment..will that be translated too?"
Yoshiyuki (Japanese):	"The appending file will not be translated. Because it is very difficult."
WonJe (Korean):	"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."

Yoshiyuki (Japanese):	" I could understand your message. There are some kinds in the TransMail's mounting.
	Type 1:
	Sender --> ServerA --> ServerB --> Receiver
	     SMTP SMTP SMTP
	TransMail is on ServerA and works as WRAPPER of SMTP.
	Type 2:
	Sender --> ServerA --> ServerB --> Receiver
	     HTTP SMTP SMTP
	TransMail works as WebApplication on ServerA. TransMail translates a message and gives SMTPD that message.
	Type 3:
	    Sender --> ServerA --> ServerB --> Receiver
	     HTTP RSYNC HTTP
	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."

* Message by Puja is posted in English. Messages by Yoshiyuki are posted in Japanese and which by WonJe is posted in Korean and translated into English by arcnet/sangenjaya.

Figure 3: Active Discussions Regardless of Translation Quality: Software Implementation Phase

procedures implemented between humans and between human and machines.

In the **Software Implementation Phase**, on the other hand, some participants began to conduct active technical and practical discussions in the context of software development regardless of translation quality (for example, see Figure 3). These active participants confirmed their collaboration results and consensuses in the submitted program source code on CVS.

4   Conclusion

In this paper, we analyzed the communication patterns on ICE2002. By observing the ICE2002, following two points were found in the collaboration process in which people share a common goal;

- Emergence of *Human adaptation to machines*: People tend to adapt themselves to machine translation capability to convey remarks properly to other team members.
- Achievement of *Human to human collaboration*: Even if a translation does not work perfectly sentence by sentence, people work out intentions of other members' remarks.

That is to say, we confirmed that machine translation indeed supported intercultural collaborative works. The approach of this research is to use translation function in collaborative context. We continue to analyze the human-to-machine interaction, especially the human recognition and adaptation process to machine translation. Then, we feed back the human-to-machine interaction model to natural language processing researchers. Our ultimate goal in this research is to develop translation-mediated groupware, which can be applied to the multinational project team activities in practical worksites.

Acknowledgement

References

Have, P. T. (1999). Doing conversation analysis: A practical guide. London: Sage Publications.

Krippendorf, K. (1981). Content analysis: An introduction to its methodology, London: Sage Publications.

Maegaard B. (Ed.). (2001). Proceedings of MT Summit VIII.

Pastore, M. (2001). Asian Internet market to surpass United States. Retrieved August 7, 2001, from http://cyberatlas.internet.com/big_picture/geographics/print/0,,5911_861481,00.html

Schegloff, E. A. (1992). Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 98, 1295-345.