

アジアにおける 異文化コラボレーション実験 2002: 機械翻訳を介したソフトウェア開発

■実験の背景

世界におけるインターネットの分布が、急速に変化している。特にアジア太平洋地域でのインターネットの急拡大は目覚しく、2003年にはアメリカ（1億6,280万人）や欧州（1億6,220万人）を追い抜き世界トップの1億8,330万人になると予測されている^{☆1}。こうしたインターネット人口の拡大にともない、インターネット上での利用言語に多様化が生じている。かつては英語を母国語とするユーザが大半だったが、2002年度にはインターネット上での英語利用者の割合は、36.5%にまで減少している。一方で、日本語、韓国語、中国語を母国語とするユーザが25.1%に上昇している^{☆2}。Webページ数は現在でも7割が英語といわれるが、この値も利用言語人口の変化に追随していくと思われる。

インターネットの世界的普及は、多様な人々とのコミュニケーションの場を拡大している。ソフトウェアのオフショア開発などビジネスの場、またLinuxに代表されるオープンソースソフトウェア開発などのボランティア活動でも、国境を越えた共同作業が低コストで可能となった。ただ、私たちが住む東アジアでは、言語の壁が、アクセスの機会平等や、相互理解促進を阻んでいる。東アジアのように、隣国の言語を教育されない国々により構成される地域にとっては、言語バリアは深刻な問題である。世界共通語とされる英語で、高度かつ論理的な議論を行うことも、東アジアの多くの人々にとっては難しい。多言語のコミュニケーションを情報技術によって支援する、機械翻訳やコラボレーション技術を駆使して言語バリアを乗り越えるという実験の構想は、こうした背景から生まれた。

インターネットを介したコミュニケーションでは、機械翻訳技術を比較的簡単に利用することができるため、技術の向上が多言語のコミュニケーションを飛躍的に拡大する可能性がある。自然言語処理などの研究分野で

京都大学大学院情報学研究科社会情報学専攻

野村 早恵子ⁱ

saeko@kuis.kyoto-u.ac.jp

京都大学大学院情報学研究科社会情報学専攻

石田 亨ⁱⁱ

ishida@i.kyoto-u.ac.jp

NTTコミュニケーション科学基礎研究所

船越 要

kf@cslab.kecl.ntt.co.jp

京都大学大学院情報学研究科社会情報学専攻

安岡 美佳ⁱⁱⁱ

mika@kuis.kyoto-u.ac.jp

NTTコミュニケーション科学基礎研究所

山下 直美

naomi@cslab.kecl.ntt.co.jp

i 2003年4月より、科学技術振興事業団CRESTデジタルシティプロジェクト研究員。

ii 科学技術振興事業団CRESTデジタルシティプロジェクト研究代表者、NTTコミュニケーション科学基礎研究所リサーチプロフェッサ。

iii 2003年4月より、東京大学先端科学技術研究センター博士課程学生。

☆1 http://cyberatlas.internet.com/big_picture/geographics/print/0,,5911_861481,00.html

☆2 Global Reach (global-resach.biz/globstats)

は、機械翻訳技術や音声翻訳技術についての綿密な研究は行われてきた¹⁾。しかしながら、研究者たちは、翻訳技術が実際のビジネスで使えるほどのクオリティに達しているとは考えていないようだ。翻訳精度への要求が厳しい新聞記事などの文書翻訳が研究対象であったためだろう。文脈依存度の大きい対話処理に機械翻訳を適用することは、さらに技術的に困難であると考えられている。自然言語処理研究者に、なぜ対話に機械翻訳技術を適用しないのかについてインタビューすると、次のような答えが返ってくる。「現在実用化されている翻訳エンジンはほぼ一文単位の処理を行っているが、意思疎通を要する場面においては、広義のコンテキスト、プラグマティクスの考慮が必要になる。また、主語や述語の省略、長文、対話や話し言葉に特徴的な表現への対処など解決しなければならない課題が数多く残っている。」

しかし、社会は、文書の正しい翻訳だけでなく、人と人を結ぶコミュニケーションの翻訳を強く求めている。また、実際、私たちのような自然言語処理研究の門外漢は、機械翻訳に人と人のコミュニケーションを支える可能性を感じる。多国籍のプロジェクトで、外国語のできる同僚がいれば嬉しいものだ。人と人のコミュニケーションは言語的に寛容ではないだろうか。コラボレーションという文脈のなかでこそ、機械翻訳の利用価値が高まるのではないか。英語を用いずに外国人とコミュニケーションができるというだけで、意思疎通時の心理的負担が軽減される。機械翻訳のミスは自分の責任ではないとの安心感もある。

この研究では、機械翻訳技術を用いる文脈を「共通の目標を持つ多言語間のコラボレーション過程」に特定し、今後の機械翻訳利用の可能性を模索する。研究のアプローチの特徴は、人々と機械翻訳システムとのインタラクションを観察し、機械翻訳技術の新たな利用を見出そうとするところにある。

■ ICE2002

《実験の準備》

言語の壁を克服するコラボレーションを目指して、京都大学社会情報学専攻、科学技術振興事業団、NTTコミュニケーション科学基礎研究所の共催で、異文化コラボレーション実験 (Intercultural Collaboration Experiment 2002 (ICE2002))^{☆3}を行った²⁾。

実験のタスクは、多国籍チームによるオープンソースソフトウェアの共同開発と設定された。ソフトウェア開発を実験のタスクとして採用した理由はいくつかある。まず、機械翻訳を、多様なテーマが話し合われるチャット

トではなく、対話の文脈を絞ることができる場に適用したかった。タスクを限定することによって、技術の改良が成果を導きやすいと考えたからである。また、共通の目標を持つコラボレーションの場に適用したかった。協調して目標達成を目指すためには、メンバー間での明確な意思疎通が必要となる。機械翻訳のミスは、何らかのかたちで克服されるに違いない。また、克服の過程を解析することにより、機械翻訳システムのメンタルモデルを明らかにできるのではと考えた。最後に、多国籍のソフトウェア開発に将来性を感じた。この背景には、アジアワイドなソフトウェア共同開発の需要の急伸がある。ソフトウェア開発には、論理的かつ綿密な議論が要求される。しかし、すべての技術者が英語に堪能なわけではない。実際に実験に参加した学生たちは、一様に「ソフトウェア開発ドキュメントを母国語で書きたい」と主張した。論文やチャットは英語で書けても、創造的な思考は母国語で行いたいのだ。また、インタビューした企業関係者は、「多国籍プロジェクトの困難さは技術的問題ではなくコミュニケーションの問題だ」と述べた。小さな意思疎通のずれが、プロジェクトの失敗を招く。ソフトウェア開発の現場で機械翻訳は必要とされるに違いない。

実験の構想がある程度固まり、次に実験参加者を募ることになった。東アジアを中心に、アジアワイドな実験を展開したいと考えた。中国と韓国は東アジアの代表国として参加してもらいたい。中国には、デジタルシティプロジェクト³⁾以来、学生交流を続けている上海交通大学副学長のシェン教授の研究室に依頼をした。韓国からは、ソウル国立大学のリー助教授が、ハンドン大学チョウ教授を誘って参加してくれた。また、マレーシアで、NTT MSC Sdn.Bhd. が、NTT 研究所開発の日英・日馬翻訳エンジンを他の言語の翻訳エンジン・サービスと組み合わせ、多言語間の翻訳サービス Arcnet/sangenjaya^{☆4}として提供していることを知った。協力を得るために、さらにマレーシアの大学からの参加者を募るため、2002年3月にマレーシアを行脚した。マラヤ大学 (UM)、マルチメディア大学 (MMU)、マレーシア科学大学 (USM) をまわった。結果として、NTT MSC Sdn.Bhd. が、日本語、韓国語、中国語、マレー語、英語の5カ国語間の翻訳サービスを提供する意向を、またマラヤ大学の CSCW 研究者が実験に参加する意向を表明してくれた。

こうして ICE2002 には、上海交通大学、ソウル国立大学、ハンドン大学、マラヤ大学、そして京都大学から総勢 40 名強の学部学生や院生 (チームメンバとチームリーダー) と教員 (チームアドバイザー) が集まること

^{☆3} <http://ice.kuis.kyoto-u.ac.jp/>

^{☆4} <http://sangenjaya.Arcnet.my/index-e.html>



図-1 ICE2002の参加者

となった。学生の実験参加動機は、「他国の学生と交流ができることや国際的なプロジェクトに関心があった」「機械翻訳を介したコラボレーションに新しさを感じた」「バーチャルに共同作業をするというコンセプトが面白い」「オープンソースソフトウェア開発に興味がある」などと積極的なものが多い。参加した学生の大半が本実験に非常に高い意欲を持っていたことは特筆に価する。図-1は、4カ国から参加した学生や教員、そして実験運営者を表している。

ICE2002は、日本からは情報処理学会（IPJS）、中国からは Shanghai Computer Society（SCS）、韓国からは Korean Intelligent Information Systems Society（KIISS）、そしてマレーシアからは IEEE Malaysia Section の協賛を得た。実験の設計として、各国の学生や教員が単なる被験者として参加するのではなく、それぞれが教育・研究を目的として参加するようにした。つまり、実験ではあるが、ある意味では実用であるように設定したのだ。実際、参加した学生の数名は、ICE2002に参加し、開発したソフトウェアで卒業論文や修士論文を執筆している。また、国内外の研究会や会議で発表をしている^{2), 4) ~ 6)}。

実験の構想、参加者を確定する作業と並行して、参加者がソフトウェア開発を行う際のコミュニケーション環境を急ピッチで整えていった。図-2に示されるように、Arcnet/sangenjaya の提供サービスでは翻訳できない言語対が残されていた。このような言語対については2段階の翻訳が試みられたが、機械翻訳を2回通すと翻訳品質が猛烈に悪くなった。そこで、日本語、中国語、韓国語の間の翻訳サービスを探すことにした。幸い、日

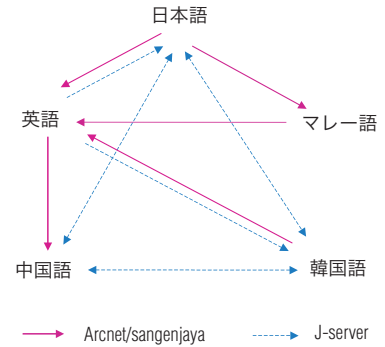


図-2 ICE2002で利用された翻訳サービス

本語と韓国語の翻訳サービスJ-serverを提供する（株）高電社^{☆5}が、大阪を拠点に事業展開をしていた。翻訳サービスの不正利用を防ぐための対策などを講ずるという条件で、研究目的としてオープンソースソフトウェア開発の場での翻訳サービスの提供を受けることとなった。こうして多言語コミュニケーションツール、TransBBS および TransWeb の準備が整った。

《多言語コミュニケーションツール》

前節に述べたように、ICE2002ではArcnet/sangenjayaとJ-serverのインターネット経由での翻訳サービスを組み込んだ多言語コミュニケーションツール、TransBBS および TransWeb を開発し、参加者に提供した（図-3 および 図-4 参照）。5カ国語の翻訳は、図-2のような組合せで行われた。直接の矢印がない言語ペアに関しては、日本語を介した2段階の翻訳が行われた。

TransBBS や TransWeb は、ソフトウェア開発プロジェクトで利用可能なほど高度なツールではない。翻訳精度はもちろんのこと、グループウェアとしての機能をほとんど有していない。私たちは、TransBBS と TransWeb を、多国籍プロジェクトにおける研究課題を洗い出すためのツールであると考えた。そのため、実験期間中においては、同国人、他国人であるにかかわらず、すべてのディスカッションをTransBBS および TransWeb 上で行うよう義務付け、対面コミュニケーション、電子メールや Chat システムなどを用いたコミュニケーションを禁止した。また、両ツールを、参加学生がソフトウェア開発をする際の教材としても位置付け、それらのソースコードを開放した。

☆5 <http://www.j-server.com/index.shtml>

《TransBBS》

TransBBS は、開発ソフトウェアの設計に関する意見交換や開発状況の報告など、実験参加者同士が日常的なやりとりを行うための多言語電子掲示板システムである。TransBBS では、ユーザ登録事項に含まれる母国語情報にあわせ、掲示板のインターフェースが5カ国語の中から自動的に選択される。たとえば、日本語を母国語とするユーザに対しては、TransBBS のインターフェースは日本語で表示される。各ユーザの選択により、TransBBS 上に投稿されたメッセージ表示言語の組合せを自由にアレンジすることもできる。また、メッセージを投稿する際には翻訳の出力結果を確認し、何度でも納得がいくまでメッセージを書き換え再翻訳できる機能などが用意されている。TransBBS 上で行われるコミュニケーションは、非対面、非同期、テキストのみであり、かつ機械翻訳を介したノイズの高いものである。このようなリーンなメディア上で円滑なコラボレーションを実現する工夫として、TransBBS に多国籍メンバで話し合うための部屋 (Lobby) と、各国で分担開発中のソフトウェアについて話し合うための部屋 (China, Japan, Korea, Malaysia Room) が用意された。

また、ネットワーク環境の発展途上の国との共同作業も見越し、TransBBS は完全分散システムとして実装された。すなわち、各国の環境にそれぞれサーバをインストールし、複数サーバ間で情報共有を図るネッ



図-3 TransBBS

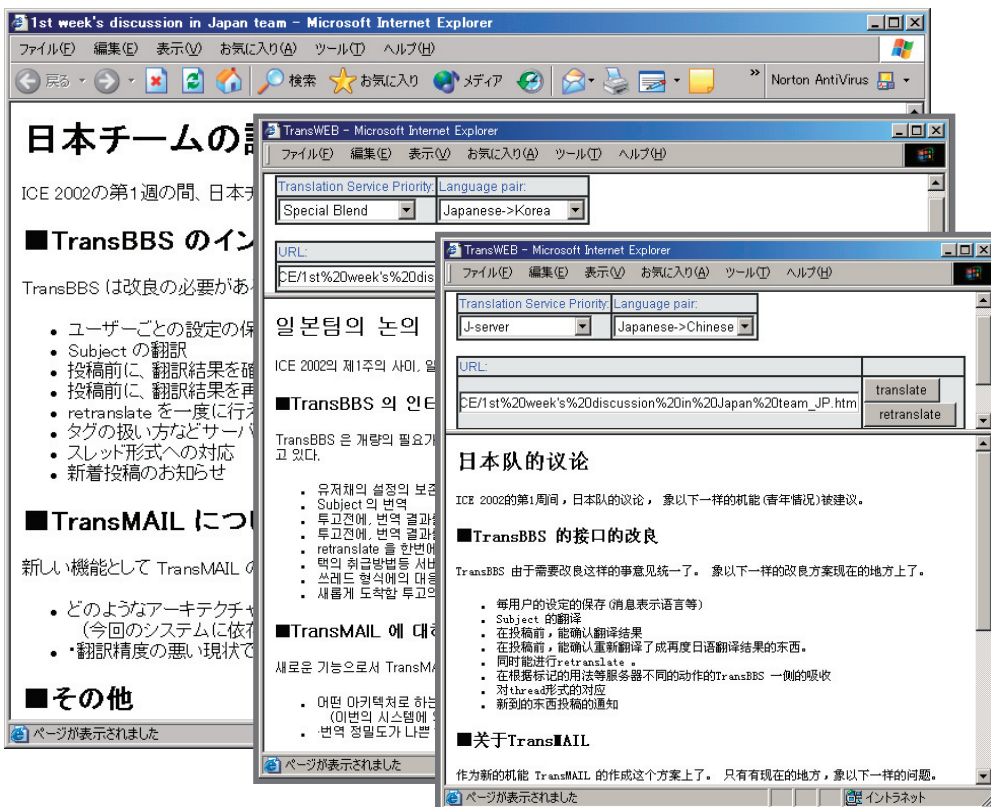


図-4 TransWeb (左：日本語原文，中：韓国語に翻訳，右：中国語に翻訳)

トワーク構成をとった (ただし、翻訳サーバは Arcnet/sangenjaya と J-server の2カ所に集中)。複数配備されたサーバ間で、常にデータの同期が図られた。そして、ネットワークが不安定な場合にはデータ変更履歴が記録され、ネットワークが回復したときに、他のサーバとデータが同期された。利用者は、ネットワークの中で最も近いサーバを用いてデータの送受信を行った。たとえば、日本サーバと中国サーバの回線が不安定になったときには、各国のサーバ内のみで投稿が継続された。そして、回線が復帰すると、サーバ間で差分データが送受信された。

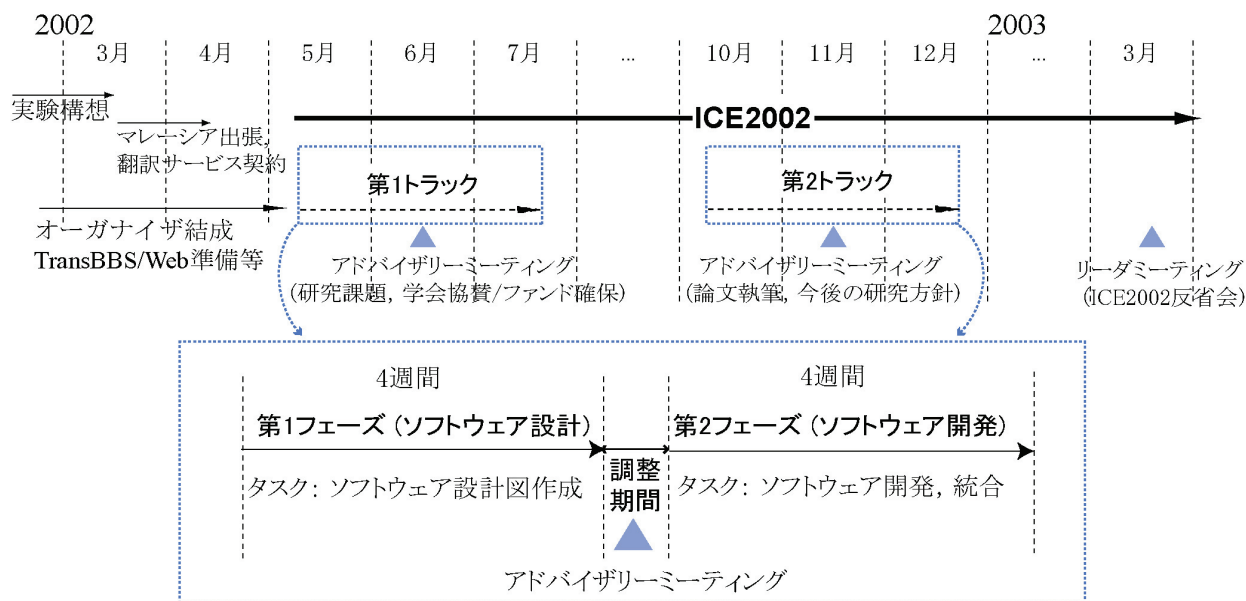


図-5 ICE2002 スケジュール

《TransWeb》

TransWebは、他国のメンバが母国語で書いたソフトウェア開発ドキュメントを読者が自分の母国語に翻訳して閲覧するアプリケーションである。本実験では開発ドキュメントはHTML形式で記述することにし、TransWebをWebページ翻訳システムとして実装した。TransBBS同様、Arcnet/sangenjayaおよびJ-serverの翻訳サービスを介し、中国語、日本語、韓国語、マレー語、英語の翻訳を実現している。

分散型プロジェクトによる効率的なソフトウェア開発には、知識共有のためのシステム構成図やデータフロー図が必要である。各国の参加者は、母国語で作成したソフトウェア開発ドキュメントをHTML形式で保存し、TransBBSを用いてドキュメントのURLを他国の参加者に伝える。それを受けて、他国の参加者は、TransWeb上でドキュメントを母国語に翻訳して参照した。

このように、ICE2002では、両ツールを利用して、ソフトウェア開発のための日常的なディスカッションとドキュメントを介した情報の共有が継続して行われた。

《実験の経過》

ICE2002は、第1トラックとして5月20日～7月19日、そして第2トラックとして10月15日～12月13日と、2度に分けて実施された(図-5参照)。これは、ICE2002が課題探索型の実験であるため、経験を通じた反復が必要だと考えたためである。

実験では、参加者は、限られた時間内に活発に情報交換を行いながらソフトウェアを完成させなければならな

かった。したがって、開発過程をスムーズに進行させるための工夫として、以下が試みられた。

1. オーガナイザチームの結成：

ICE2002を運営するためのオーガナイザチームを文系理系の学生・若手研究者で結成した。オーガナイザは、実験前にはTransBBSやTransWebの動作テスト、翻訳の質の確認と翻訳サービスの選定、ICE2002のWebサイトの構築、TransBBSやTransWebの配布、各国のアドバイザーやチームリーダーとの連絡、実験運営に関するマニュアルや機械翻訳を上手に利用するためのノウハウ集(図-6参照)などを作成した。

実験時には、TransBBS上でのコミュニケーションを円滑にするための“Overkind Coordinator,” 技術的なアドバイスをする“Technical Supporter”, 投稿文章の文法チェックをする“Translation Guerrilla”などの役割を担い、参加者を支援し刺激し続けた。さらに実験期間を通じて、参加者全員に対して週報(weekly summary)を流した。週報では、コミュニケーションの状態、各国での開発状況の進捗などを報告した。実験終了後には、実験運営や翻訳の質などに関するアンケート調査を行った。

オーガナイザチームが文理融合であったことが、成功の鍵となった。文系学生の熱心なオーガナイジングと、理系技術者の専門的知識が、参加学生の活動を支えた。

2. 実験期間の分割とソフトウェア開発の方法：

- ・実験の各トラックは4週間ずつの「ソフトウェア設計フェーズ」と「ソフトウェア開発フェーズ」に

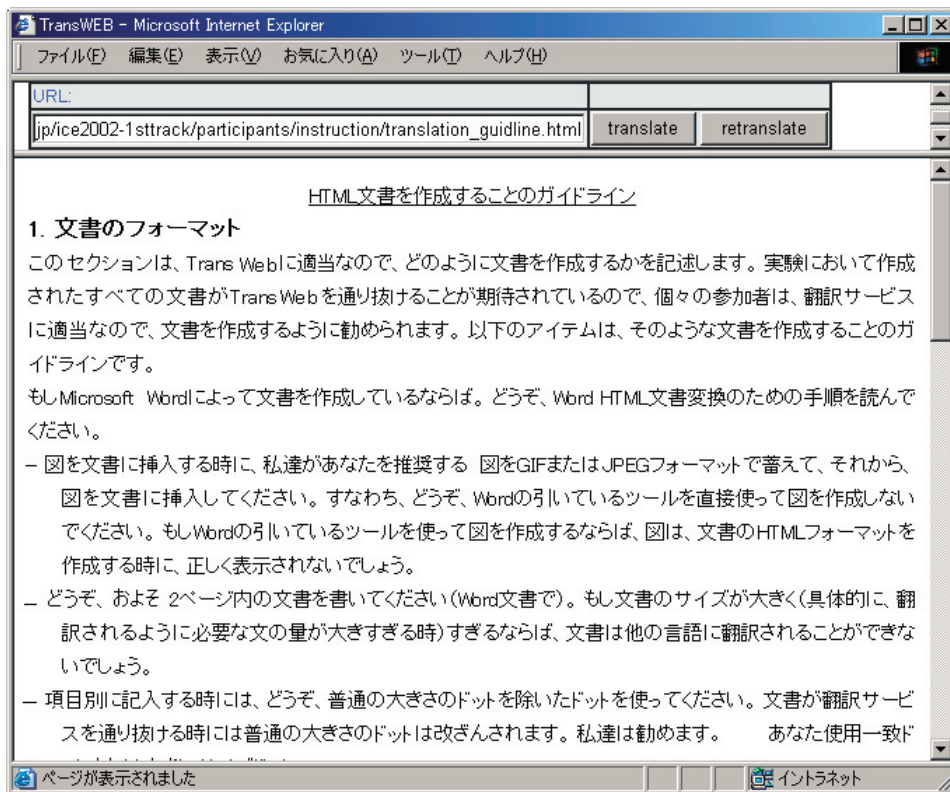


図-6 翻訳利用のためのガイドライン（英語を TransWeb で日本語に翻訳）

分けるかたちで構成された。それぞれ週単位である程度タスクが決められた。また、ソフトウェア設計フェーズとソフトウェア開発フェーズの間に2週間の調整期間を設けた。その間に、各国のアドバイザーとオーガナイザが対面で（京都でアドバイザーミーティングを行った）集まり、その後の実験について、また各国での研究の進め方などについて意見を交換した。「ソフトウェア設計フェーズ」と「ソフトウェア開発フェーズ」とに分けたのは、両者でコラボレーションの性質が異なると思ったからである。前者では発想支援や合意形成型のコラボレーションが、後者ではLinuxの開発のように行動主導型のコラボレーションが行われると想定した。

- ソフトウェアのモジュール設計および開発は、基本的に国ごとに分かれて行われた。ただし、最終的には整合性のある単一インタフェースのソフトウェアに仕上げて提出するという制約を課した。このため週1度のペースで、各国が開発ドキュメントを作成し、それらを用いて情報を共有し、議論を行うというルールが設定された。作成されるソフトウェアの仕様は、UMLを用いて記述された。国ごとに分かれて開発したのは、それが最も自然な形態であると考えられたためである。実際問題として、ICE2002を純粋な統制実験とすることは不可能だった。多数の教員・学生を1年間拘束するためには、各国チ

モジュール	機能
TransSearch (中国)	ユーザにより入力された質問語を5カ国語のうちユーザの選択にあわせて翻訳し、その質問語を、Googleを用いて検索する。Googleの結果を5カ国語に翻訳するサービス。
TransMail (日本)	5カ国語翻訳機能を組み込んだWWWベース電子メールサービス。
TransChat (韓国)	5カ国語翻訳機能を組み込んだチャットサービス。
TransSMS (マレーシア)	主にPDAやモバイル機器を対象にした160文字までのショートメッセージサービス。長文がうまく訳されないという教訓から開発されたサービス。

表-1 開発されたソフトウェアモジュール

ームがそれぞれ自主的に研究テーマを設定し、活動することが必須と思われた。その結果、各国チームはモチベーションを維持し、開発したソフトウェアを題材として修士論文や国際会議論文を執筆している。一方で独立した開発が、特にソフトウェア開発フェーズでの対話を減少させたことは否めない。

- ソフトウェア開発フェーズでは、ソフトウェア開発のためのバージョン管理システム(CVS)が運営サーバに開設され、分散環境におけるソフトウェア開発の効率化がはかられた。また、各国が開発したモジュールを統合する役割を持つソフトウェア統合チームが構成され、それぞれのモジュールを最終的に1つのインタフェースに統合した。ソフトウェア統合チームは、中国、日本、韓国、マレーシアの各国

の参加者から1名ずつが選出され、結成された。

各国チームが開発したソフトウェアモジュールを表-1に示す。結果的に本実験で各国が開発したモジュールは、すべて機械翻訳を組み込んだものとなった。TransXXという命名も、各国の自立的な判断によるものである。

■実験から得られた知見

TransBBS上のユーザアクションログ（メッセージ翻訳、投稿、閲覧、など）は、実験期間全体で31,000件に及ぶ。ソフトウェア開発の現場で、参加者はどのような工夫をしながら、機械翻訳というノイズの高いコミュニケーションチャネルを利用したのだろうか。本章では、TransBBS上でのユーザアクションログとメッセージログを用いてコミュニケーションパターンを分析する。

《翻訳システムへの利用者適応》

TransBBSに投稿されるメッセージを見る限りでは、機械翻訳が提供する翻訳品質は、お世辞にも高いとはいえない。実験の初期には、挨拶と同時に翻訳品質に関するメッセージが目立った。「他国の参加者が投稿したメッセージの意味が分からない」という不満が投稿された。自分の投稿メッセージに対する翻訳誤りを修正し、他国の参加者とコミュニケーションを図ろうとする行為も観察された。たとえば、次のメッセージは、自国のメンバに対するメッセージでも、多国の参加者に理解してもらうためには、何に気をつけるべきかを考察している。

「日本のDiscussionルームでは、どの程度翻訳精度を気にかけるべきですか？日本語で通じてしまうと思うと、思わずそのまま投稿してもよいのではと感じてしまうのですが、他のチームから何を議論しているのか分からないのは問題ですね。気をつけるようにします。（原文：日本語）」

ノイズの高いコミュニケーションチャネルを克服し、相手に正しく意図を伝えるために行った参加者の行為は以下の2種に分類できる。本稿では、こうした翻訳誤りを解消する行為を、従来の会話分析研究に習い「リペア」と呼ぶ。

- ー 自己主導型リペア (Self-initiated repair) : TransBBSに投稿する前に、翻訳品質を上げるために何度も原文の精練を自主的に行う行為。多くの場合、唯一の共通言語である、英語に翻訳された結果を手がかりに行われる。なお、英語の訳文を頼りにしてリペアが行われると、それ以外の言語（たとえば、韓国語が原文の場合は、日本語、中国語、マレー語）でも明らかに翻訳の質が上がるのが、各国からの参加者の協力により確認されている。
- ー 他者主導型リペア (Other-initiated repair) : メッセー

ソフトウェア設計フェーズ	
参加者名	投稿回数/ 投稿前翻訳回数
Saeko	69/234
Bikesh	66/112
Tomoko	51/395
Yoshiyuki	35/170
LimSanny	28/71

表-2 TransBBSでの自己主導型リペア

japana-149 || 投稿者: FUJISHIRO Yoshiyuki || document
Subject: We want to know whether you got my message
投稿日時: 2002年05月23日 木曜日 13時58分42秒

日本語	英語(retranslate)	中国語(retranslate)	韓国語(retranslate)
あなたが私の投稿(japana-148)を理解できたかどうかを私たちは知りた。あなたが理解した部分をおあなたの言語を使って投稿してください。また、(japana-148)に対するあなたの意見も投稿して欲しい。	We want to know whether you could understand the my contribution japana-148 or not. We, please contribute using your language the part that you understood. I want us to also contribute you to japana-148's opinion.	你能理解与否我的投稿(japana-148)我们想知道。你理解的部分使用你的语言请投稿。同时,对(japana-148)的你的意见也请投稿。	당신이 나의 투고(japana-148)을 이해할 수 있었던 것인가 아닌가를 우리들은 알고 싶다. 당신이 이해한 부분을 당신의 언어를 써서 투고해 주십시오. 또한, (japana-148)에 대한 당신의 의견도 투고해주었으면 싶다.

図-7 機械翻訳に適応した文章作成（赤字枠は投稿言語）

ジの受信者が、意味の確認や、意味が分からないというリアクションを行うことにより、翻訳の精練作業がメッセージ投稿者と受信者の間でインタラクティブに行われるもの。

TransBBSに投稿するメッセージに、自己主導型リペアが何回行われたかを表-2に示す。たとえば、Tomokoはソフトウェア設計フェーズで、平均すると1つの投稿メッセージにつき約8回のリペアを行っている。また、図-7に示すメッセージには、翻訳誤りを減らす工夫が各所に現れている。主語を徹底的に補った文章から、参加者の翻訳システムへの涙ぐましい適応への努力を見てとることができる。参加者はTransBBSを使っていくうちに、翻訳に向かない文章構造（主語や目的語を省くこと）が誤った翻訳結果を生み出すことを理解し、それを避けるために翻訳サービスに適応した文章を用いたのである。なお、自己主導型リペアの頻度は、実験が進むにつれ減少した。たとえば、ソフトウェア設計フェーズにおける平均リペア頻度が4.97回であったのが、ソフトウェア開発フェーズでは2.04回になった。

また、他者主導型のリペアでは多国間で翻訳の意味確認が行われた。たとえば、日本人の参加者が挨拶の場で「よろしくお願ひします」という日本語独特の挨拶文を投稿したことがあった。「よろしくお願ひします」は、会話のコンテキストにより意味が異なる。TransBBS上

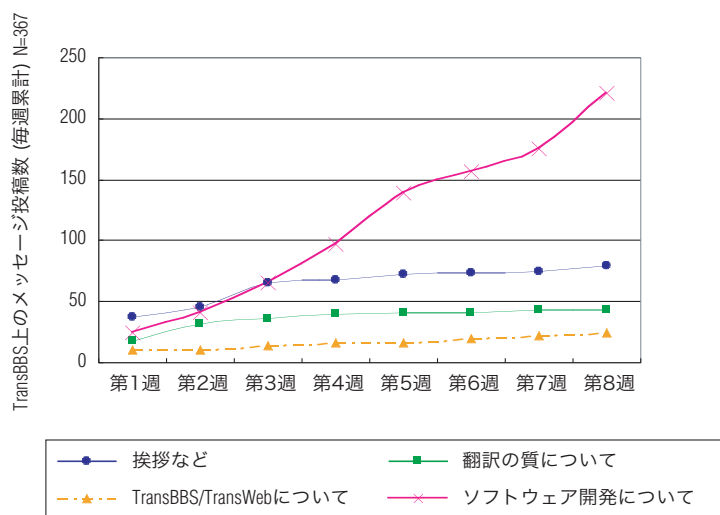


図-8 実験過程における TransBBS 上でのトピックの推移 (5月20日～7月19日)

で、誤って訳された「よろしくお願いします」の本来の意味を確認するやりとりが(多言語で!)行われ、最終的にメッセージの正しい理解に到達している。

《翻訳の協調的精練》

図-8は、第1トラック(5月20日～7月19日まで)における、TransBBS上での議論のトピックの推移を表している。横軸は時間(週単位)を表し、縦軸は投稿されたメッセージ数を表している。実験が進むにつれ、翻訳品質に対する意見はほとんど発せられなくなった。同時に、挨拶やインフォーマルな発言も観察されなくなった。その一方で、ソフトウェア開発に関する意見交換が増えていった。これは、参加メンバが、翻訳結果は完全ではなくても対話を通じて意味を確認し、応答を与えていたことによる。たとえば、図-9は日本チームが中心になって開発したTransMailに対する他国メンバからのコメントである。TransMailはこのようにして、さまざまな参加者からの意見を反映し完成された。

ICE2002のコミュニケーションパターンは、次のようにまとめることができる。実験を通じて翻訳サービスの品質は変わらなかったのにもかかわらず、参加者は翻訳サービスに適応し、その結果、翻訳リペアの頻度は減少していった。さらに、実験が進むにつれ、議論されるトピックは本来の目的追及のためのものに収束し、積極的な議論が行われた。

■実験の今後

ICE2002を通じて明らかになった点は、以下のとおりである。

- ・共通の目標を持った者同士の間であれば、人はたとえ

自然言語処理の専門知識がなくとも機械翻訳とのインタラクションを通じて自らを翻訳機能に適応させ、言語を越えて理解を深めるよう努力する。

- ・特定の知識を共有する者同士のコラボレーション過程であれば、翻訳の誤りは比較的容易に克服され、コミュニケーションを成り立たせることができる。

ICE2002を通じて私たちは、機械翻訳は人々の異文化コラボレーションを支え得る技術であるとの確信を持ち始めている。つまり、ソフトウェア開発などのように目標を絞ったコラボレーション過程に対して、機械翻訳は適用可能であると考えている。研究者は完全な技術を追求するが、利用者は有用な機能を必要としている。不完全な機械翻訳技術を、国境を越えたプロジェクトに役立たせる知恵は、文系・理系のバックグラウンドを持つ研究者・利用者の協同作業によってのみ可能だと思う。ICE2002の特徴は、研究グループに社会科学と計算機科学をバックグラウンドとした研究者・学生が参加していたことである。

多くの研究課題が残されている。現在、ICE2002のログデータを用い、プロジェクトの各段階での翻訳リペアのタイミングや頻度、質疑応答のタイミングや議論の収束過程をより詳細に分析している。今後は、社会科学的手法によるコミュニケーションの分析結果を、システムの設計開発にフィードバックさせたい。また、自然言語研究者との共同研究を進めていきたい。

興味深い例として、自己主導型リペアを盛んに行った参加者は、機械翻訳の機能を理解することの難しさを以下のように述べている。

「普通の会話であれば障害なく即座に相手に伝えることができるものを、機械翻訳を介していることによって、何度も自分の伝えたい内容を変化させ、吟味し、時間をかけてコミュニケーションを行っていたと感じます。何度精練を試みても翻訳がうまくいかず、不本意ながらそのまま投稿したり、投稿そのものをあきらめたこともありました。本当に重要なことに関しては、どうあっても翻訳を精練させて相手に伝えようとするのですが、その他の細かい内容や微妙なニュアンスは、多くの場合投稿時には失われていたと感じます。」

「普通の手紙であれば障害なく即座に相手に伝えることができるものを、機械翻訳を介していることによって、何度も自分の伝えたい内容を変化させ、吟味し、時間をかけてコミュニケーションを行っていたと感じます。何度精練を試みても翻訳がうまくいかず、不本意ながらそのまま投稿したり、投稿そのものをあきらめたこともありました。本当に重要なことに関しては、どうあっても翻訳を精練させて相手に伝えようとするのですが、その他の細かい内容や微妙なニュアンスは、多くの場合投稿時には失われていたと感じます。」

機械翻訳はチームの中でどのような存在であるべきなのだろうか？ 私たちは現在、多くのリペアログを基に、人が機械翻訳に期待する翻訳結果と、実際の翻訳結果とのギャップを分析している。機械翻訳が人々のコラボレーションの文脈で用いられた経験は多くない。機械翻訳研究者は、システムのメンタルモデルを意識してきただ

malay-252 || 投稿者: Puja Lakshmi || [document](#)

Subject:
投稿日時: 2002年06月03日 月曜日 13時39分33秒



日本語(retranslate)	英語	中国語(retranslate)	韓国語(retranslate)	Malay(retranslate)
メール横断はよいアイデアです。メールが持っていることおよびなくなくてかし attachment .will that 翻訳 また ?	Transmail is a good idea. But how will the content be translated and if the mail has an attachment .will that be translated too?	Transmail是一个好主意。但是如何将内容被翻译和如果邮件有一份附件。。也被翻译的意愿?	메일(mail) 횡단은, attachment . will that 번역 도이어서, 만약 메일(mail)의 가장 의미내 용번역의에 대해서, 미처 갖게 하고 있는 좋은 idea. But입니까?	Transmail adalah suatu cadangan baik. Tetapi diterjemahkan bagaimana akan kandungan dan jika surat mempunyai suatu lampiran .. Akan yang diterjemahkan juga?



japana-254 || 投稿者: FUJISHIRO Yoshiyuki || [document](#)

Subject:
投稿日時: 2002年06月03日 月曜日 19時53分43秒



日本語	英語(retranslate)	中国語(retranslate)	韓国語(retranslate)	Malay(retranslate)
Re: malay-252. 添付ファイルは翻訳されないでしょう。なぜなら、それは非常に難しいからです。	Re : malay-252. The appending file will not be translated. Because it is very difficult.	Re:malay-252. 附加文件不被翻译吧。要是为何, 那非常难。	Re: malay-252. 첨부 파일은 번역되지 않지요 왜냐하면, 그것은 대단히 어렵기 때문입니다.	Malay-252 : Re. Fail appending akan ditranslate. Because itu cukup difficult.



japana-272 || 投稿者: Won Je Yang || [document](#)

Subject:
投稿日時: 2002年06月05日 水曜日 16時02分57秒



日本語(retranslate)	英語(retranslate)	中国語(retranslate)	韓国語	Malay(retranslate)
TransMailの場合 貼付ファイルの問題は そのように(そんなに)重要に見えるか しいですね。SMTP を経るに Wrapper 役割をして上げる server を経れば十分に可能だろう 見ます。	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.	TransMail机动警察右粘贴文件疑问制那样(那样)重要河事是波特不有。经过SMTP前粗暴的多面充分地可能下更短发啦Wrapper作用乙亥蕉耳server.	TransMail의 경우 첨부 파일의 문제는 그렇게 중요하게 보이지 않는군요. SMTP를 거치기 전에 Wrapper 역할을 해줄 server를 거친다면 충분히 가능하리라 봅니다.	Anda dapat dilihat atau tiada suatu masalah fail lekat tersebut dalam hal TransMail secara penting sukakan ini lebih demikian) . Ia melihat yang akan menjadi secara cukup munasabah jika melalui pelayan untuk mengangkat dengan bermain peranan Balutan melalui SMTP

図-9 翻訳の質を越えたコラボレーション (赤枠内は投稿言語)

ろうか。コラボレーションの中での機械翻訳は、人工知能とヒューマンインタフェースに新しい研究の題材を提供するように思う。

また、異なる国々での文化的差異も研究対象として考えていきたい。ソフトウェア開発プロジェクトにおいても、スケジュール管理やミーティングの進め方などに相違があるだろう。言葉遣いや他者への対応も文化により左右される。アジアでのソフトウェア共同開発は、技術的課題以上に言語・文化の違いがプロジェクト遂行に支障をきたすケースが多いと聞く。実際のソフトウェア開発の現場に入り、ビデオを用いた観察やインタビューを行い、比較検討することを考えている。

ICE は技術的シーズから出発した研究プロジェクトではない。ICE は 2005 年を目途に、実際の多国籍プロジェクトへの異文化コラボレーション環境の提供を目的としている。そのための、コラボレーション過程の分析とシステム開発を並行して進める。また、アジア各国の大学・企業との交流を通じて、東アジアにおけるコラボレーションに技術的・社会的に貢献できたらと考えている。

謝辞 NTT MSC Sdn.Bhd. (Arcnet/sangenjaya) の小寺博氏, 阿部明典氏, (株)高電社, NTT コミュニケーション科学基礎研究所の大山芳史氏, 佐藤哲司氏, 真鍋義文氏, 桑原和宏氏, NTT サイバースペース研究所の林良彦氏, 小倉健太郎氏, 松尾義博氏, (株)数理システムの山本晃成氏, 京都大学の坂本知子氏, 藤代祥之氏, 岡本昌之氏, 中塚康介氏から多大なるご協力をいただいたことを深く感謝いたします。

参考文献

- 1) Maegaard, B. Ed.: Proceedings of MT Summit VIII (2001).
- 2) Nomura, S., Ishida, T., Yasuoka, M., Yamashita, N. and Funakoshi, K.: Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002, HCII2003 (2003).
- 3) Ishida, T.: Digital City Kyoto: Social Information Infrastructure for Everyday Life, CACM, Vol.45, No.7, pp.76-81 (2002).
- 4) Funakoshi, K., Yamamoto, A., Nomura, S. and Ishida, T.: Lessons Learned from Multilingual Collaboration in Global Virtual Teams, HCII2003 (2003).
- 5) Othman, M. and Lakhmichand, B.: TransSMS: A Multi-Lingual SMS Tool, HCII2003 (2003).
- 6) 小倉健太郎, 林 良彦, 野村早恵子, 石田 亨: 目的指向の異言語間コミュニケーションにおける機械翻訳の有効性の分析-異文化コラボレーション ICE2002 実証実験から-, 情報処理学会第 65 回全国大会, 2T6-4 (2003).

(平成 15 年 3 月 4 日受付)

