

AAAMT The Asia-Pacific Association for Machine Translation Journal



アジア太平洋機械翻訳協会

The Fifth International Conference of Theoretical and Methodogical Issues in Machine Translation $\begin{bmatrix} T & M & I & '9 & 3 \end{bmatrix}$

The Fifth International Conference of Theoretical and Methodogical Issues in Machine Translation

〈第5回機械翻訳の理論的及び方法論的問題に関する国際会議〉

メインテーマ次世代の機械翻訳開催期間平成5年7月14日~16日開催場所京都国際交流会館

プログラム

· 自然言語処理技術(文脈処理、Non-Literalな文章等)

(概要)

- ・ 機械翻訳の新しいアーキテクチャー
- · 多言語間機械翻訳
- ・ 音声処理技術との結合
- ・ 機械翻訳用大規模知識ベース
- ・ 機械翻訳システムの評価
- ・ 機械翻訳システムを使いやすくするための技術
- ◆ 投稿論文は2月末日で締め切りです。

参加登録料

20,000円 5月末までに登録、入金済みのもの 30,000円 6月以降登録入金のもの 参加登録料にはプロシーディングスが含まれます。

主 催 自然言語処理学会 (Association for Natural Language Processing) 協 賛 アジア太平洋機械翻訳協会(AAMT) 国際機械翻訳連盟(IAMT)

会長:田中穂積(東工大)・プログラム委員長:松本裕治(京大)・委員K.Church・平川秀樹・飯田仁・P.Isabelle・北野宏明・丸山宏・野村浩郷・S.Pulman・C.Rohrer・杉村領一・徳永健伸・J.Tsuji・M.Vasconcellos・Y.Wilks

事務局・〒141東京都品川区西五反田7-1-9 HSビル(株)アイ・エヌ・エス内「TMI'93」事務局☎ 03-3494-1869

● 目 次 ●

エッセイ「再びMTサミットが日本で」	(1)	新製品の紹介 シャープ	(4)
「翻訳フエア'93」開催報告	(2)	新製品の紹介 富士通	(6)
研究機関紹介「沖電気工業」	(13)	新製品の紹介 ロゴヴィスタ	(8)
翻訳の現場から「機械翻訳時代の翻訳者の役割は?」	(21)	新製品の紹介 CSK	(10)
ユーザ事例紹介	(14)	現行製品一覧	(11)
MT技術早わかり「文生成技術」	(16)	トピックス「業界の話題」	(22)
研究会報告「システム評価研究会」	(12)	広告 「TM I '93」	表2
海外研究動向「MT評価ワークショップ」	(20)	広告 「MT-Summit IV	表3

『再び、MTサミットが日本で』

アジア太平洋機械翻訳協会 (AAMT) が昨年6 月正式に発足してから新たな飛躍を目指す初めての 新年を迎えられました。日本機械翻訳協会 (JAM T) がその設立後わずか1年余りで、アジア諸国か らの大きな期待のもとに拡大されたことは、長尾 J AMT会長を中心とした精力的な御尽力によるもの です。

本年は、AAMTにとっては大変重要な年を迎えられました。国際的な面でのビッグ・イベントである第4回機械翻訳サミット(MT Summit IV)が(社)日本電子工業振興協会とAAMTの共催で、7月20日~22日に神戸で開催されることになっているからです。MT Summit I が1987年箱根で開催されてから6年ぶりの日本での開催であり、MT Summit 組織委員会(長尾AAMT会長)を中心に活動を開始しております。私はこのMT Summit IVの展示委員会委員長という重責を任命されました。そこで成功を目指して精一杯頑張りたいと思います。

MT Summit Iが開催された1987年当時もそうであったように、いまでも機械翻訳システムの研究開発には多くの企業や研究機関等が相当数の優秀な人材と多額の資金を投入し、実用化を目指した商用システムの開発を続けてきています。しかしながら、ユーザ側としてはその実用性についてはまだまだ懐疑的な目で機械翻訳システムを見ており、各メーカは商用システムの市場開拓・拡大に苦戦しているのが現状ではないかと思います。

当初、 MT Summit は以下のようなことを メーカ、ユーザ、政府などに広く認識してもらうこ とが必要であるという主旨で開催されたと聞きおよ んでおります。

- ・機械翻訳システムの現状の技術レベルとこれ からの解決すべき技術的問題を明らかにする。
- ・現状をユーザに十分認識してもらい、使用範囲を限定したり、運用方法を工夫することにより、コストパフォーマンスの点で人間翻訳より十分な効果が発揮できる。
- ・現状の商用システムにも、まだまだ言語処理

日本電気(株) 取締役支配人 吉川 英一

の立場から研究開発しなければならないこと が残積しており、これらの解決には長時間か かるため政府当局その他からの研究支援が必 要である。

MT Summit IVでもこの当初の目的に対する認識を更に深めてもらいたいと思います。

つまり、機械翻訳市場が健全に育成され、将来にわたりよりよい研究開発計画として継続し発展していくように、日本を始めとしたAAMT参加国の先端技術を南北アメリカ地域機械翻訳協会(AMTA)およびヨーロッパ地域機械翻訳協会(EAMT)に参画している欧米各国に十分認識してもらいたく思います。そして、各国においても機械翻訳の研究開発や商用化が途切れることなく盛んになるような契機を与えられるように MTSummit IV を成功させることが重要であると思います。

また、AAMTの活動成果の一つとして、機械翻 訳システムの実用化には徐々にではありますが翻訳 業界を先導として普及の気運が芽生え始めてきてい ると思います。このことは、昨年12月に(社)日 本翻訳協会と共催で開催された翻訳フェア '93に おける「ユーザのためのMT活用セミナー」で、先 進ユーザでのステップ別機械翻訳導入・活用方法/ こうして使っている機械翻訳/ネットワークによる 翻訳の可能性等についての報告にもありましたし、 今後も機械翻訳システムの社会浸透のためにAAM Tが果たされる役割はますます大きいのではないか と思います。そして、メーカの立場からすれば多く のユーザ事例が関係者の大きな注目を集める場で報 告されることによって、もっともっとユーザ層が拡 がりビジネスの芽が少しでも出てくるでしょうし、 それによって機械翻訳システムの研究開発は地道で はあるが着実に進展していくことと思います。

機械翻訳の理想像は21世紀を目指しています。 その目的のために技術的にクリアすべき要素技術を 明らかにし、かつ機械翻訳の果たす役割を追求する ためにも今後のAAMTの活動に大きく期待してお ります。



「翻訳フエア'93| 開催報告

平成4年12月15日~17日、東京北の丸公園内科学技 術館において通産、外務、労働、科学技術等各省庁 の後援の下に「翻訳フエア'93」を開催した。

メインテーマは「電子コミュニケーションの未来」であるが、翻訳を電子化するツールとして機械翻訳システムを位置付け、その有効活用の方法を解明する手段を論議し、また研究開発の最先端の情報を提供し、コミュニケーション環境の未来像を探ることを目的として「翻訳フエア」を企画した。

この「翻訳フエア」には所謂川上の研究開発部門から技術開発、供給、普及、川下の最終ユーザに至るまで翻訳にかかわるすべての階層の人々に参加いただくために(社)日本翻訳協会と共催、(財) 国際情報化協力センターの特別協賛という実行組織体とした。

(社)日本翻訳協会 (JTA) は労働省所管の法人で、 主たる事業は「翻訳士の資格認定試験」の実施で あり、また多くの翻訳者の育成に努めている団体 でもある。

(財) 国際情報化協力センター (CICC) は通産 省の所管で発展途上国に対するコンピュータリゼー ションへの協力、普及啓蒙活動を行っているが、1987 年より国の委託をうけて中国、タイ、マレーシア、イ ンドネシアの4カ国と最先端の多言語機械翻訳シス テムの共同研究を実施している団体でもある。

「翻訳フエア」の開会式(12月15日)には来賓、主催者側企業の代表や関係者が参集し、通産、労働両省の幹部からのご祝辞を賜った後来賓代表、長尾大会会長、湯浅実行委員長等の手によりテープに鋏が入れられ華々しく開会の幕が切っておとされた。

(12/15 シンポジウム1)

多言語機械翻訳システム

シンポジウム1は3つのセッションで構成された。 先ず東工大の田中教授から「国際化の中での機械 翻訳研究の現状と課題」と題して基調講演が行われた。

世界各国の過去から現在に至る機械翻訳の研究開発の動向が紹介され、ついでCICCが取組んでいる研究が現在市販されているトランスファ方式とは異なり中間言語方式を採用しているとの紹介があった。これは多言語間翻訳を行う場合には開発がしやすいし、また人工知能の色々な技術がこれに集約

されていて、この技術を各国に根付かせ、また技術 の波及効果も大きいという点から採用された。現実 に各国が独自にシステムを作ることが出来たし、作 られた辞書も単に機械翻訳だけではなしに自然言語 処理を理解するなどの側面からも大きな財産となっ ていると思う。今後の課題としてはやはり大規模な 知識ベースの構築、ヒューマンインターフェース、 評価法、標準化、ネットワーク、並列処理などの技 術との融合も進めて行くことの重要性が述べれた。

次いでCICCプロジェクトの概要の説明が行われた後共同研究先のタイ、中国、マレーシア、インドネシア各国から「研究開発の現状と問題点」が報告された。各国のおかれている立場や環境の違いはあるもの、何とかこれを成功させたいという熱気が感じ取れた。

最後に各国研究機関の代表者による「多言語機械 翻訳に向けての国際協力」というテーマでパネルデ スカッションが展開された。

(12/16 セミナー)

ユーザのためのMT活用法

第1セッション「ステップ別機械翻訳の導入、活用のノウハウ」では牧野助教授(東邦大)から機械翻訳をエンジニアリングとしてとらえプロセス、品質、翻訳技能の向上策が論ぜられた。小川氏(日経印刷)は「MTがそのままで使えるという人は勿論翻訳の素人だが、使えないと云う人はまたMTの素人だ」と前置きし辞書構築のノウハウを披瀝、また機械翻訳業の桜井氏は「プロでも完全な翻訳は出来ない」とし、日本文の編集の仕方を、森氏(バベル)は人材教育の現場からMT活用のポイントを講演して戴いた。

第2セッション「こうして使っている機械翻訳」では原田氏 (IBM)が5年間の使用経験を通じ、運用面、設計面における改善点やプロセスを詳細に説明、活用にあたっては翻訳の流れと特徴を把握しこれにあわせて機械翻訳プログラム、周辺プログラム、運用態勢を整えることが大切であると強調された。安藤氏 (+印) はMTはちょっと工夫すれば活用出来るとして使い方の実例を紹介、小沢氏 (日本NCR)は機械翻訳は下訳として使用しているがこれは有能な翻訳者と機械システム調和と共存という長期的な

テーマに基づいての対応である。 翻訳者がMTシステムを使いこなせるようになれば仕事の場所的制約から解放され、頭脳を単なる肉体として用いる類の翻訳の苦行から解放されるであろうと論じた。

古庄氏(翻訳工房夢屋)は会員60名を擁する翻訳を 主題としたパソコン通信ホスト局を運営しているが 聴訳式英語学習教材作り等の事例紹介があった。

第3セッション「ネットワークによる翻訳の可能性」では井上氏(イリス)がNIFTY-SERVEを使用してヨーロッパ言語と英語の機械翻訳サービスの実施状況を、原田氏(高電社)からは日韓翻訳システムをパソコンネットを利用しての使用事例の紹介があった。柏原氏(スバル)からは「機械翻訳ー商用データベース」と「機械翻訳ー契約翻訳者」というネットワークサービスの事例が、また石崎教授(慶大)からは大学内におけるゼミや授業に使用されているMTシステムの使用事例と、学生たちは「自動翻訳」と「機械翻訳」を混同していたこと、システム作りの難しさを痛感したなどの感想の紹介があった。

第4セッションは「機械翻訳と著作権、知的財産 権」と題して吉田弁護士の講演があった。

(12/17 シンポジウム2)

電子コミュニケーションの未来

第1セッション「電子コミュニケーションの未来」 西垣教授(明大)から21世紀にむけての国際コミュニケーションは電子メディアによってサポートされて行く可能性が極めて高くその中でも機械翻訳は中心的な役割を果たすとの基調講演があった。

第2セッション「機械翻訳技術の未来」と題し、中 堅研究者によるパネル討論がおこなわれた。機械翻 訳システムが完全にクリアしなければならない解析 技術やパターン認識技術の活用、人間翻訳者がもっ ているノウハウをどのように組み込むか、翻訳者と 機械翻訳とのギャップを埋め、ツールをどう埋め 込んで行くか、翻訳効果を高めるための自動前処理 技術の確立の必要性、計算機にどのように常識を入 れ込めば良いのか、またその常識の範囲などについ て熱烈な討論が展開された。

第3セッション「翻訳の未来像」では理系、文系の学者によって、機械翻訳の出現により翻訳と云う仕事が将来どのように変化して行くのか文化的側面からパネル討論が展開された。

技術固有の問題のために言語や伝統的な文化がゆがめられはしないか、機械翻訳システムでは感情が発揮できるのか、機械に強制されて無味乾燥な文章を書くようにならないか、技術が導入される初期の段階では色々な危惧が出るし、それもはづれることが多いが技術の進歩は人間や環境を変化させ、新しい文化を築いて行く事だけは真実であろうとの大所高所からのユニークな討論が展開された。

(12/15、16機械翻訳展示会)

12/15~16日に開催した機械翻訳システム及びシステム関連ソフト、MT関連ツール展示会には3団体11企業が出展した。

(財)国際情報化協力センター(CICC)ブースでは 国際共同研究先のタイ、マレーシア、インドネシア、 中国の研究員も参加し、過去6年間の研究成果を公 表する実証評価試験のデモンストレーションが行わ れた。この機械翻訳では同時多言語間の翻訳ができ る。現在5カ国間で20通りの翻訳が可能である。ま た国際ネット回線を利用すれば ASEAN 会議も日本 に居ながらにして各国代表の演説を知ることも可能 という画期的システムでもある。まだ研究開発段階 にはあるが、早急な商品化が大いに期待される。

海外生産シフトの展開、海外研究開発拠点の設置、 国際情報交流活動の活発化など国際化グローバル 化は著しく進展しており、それだけに翻訳ニーズも 多方面に拡大しているが、 今回の展示会開催に当た っては研究産業、輸出入産業、海外生産比率の高い 電機、情報、事務機、自動車などの産業界にターゲ ットをしばり重点的に来場促進のDMやPRを行った 結果2日間で3500名のお客様をお迎えすることが出 来た。会期中は来客が途切れる事もなく、とくに翻 訳ニーズの高い部門のお客様だけに滞留時間も長く、 熱心な質疑応答が繰り返された。特に会期直前に新 製品を発表した数社のブースでは黒山の人だかりの 活況を呈し、説明員は本当に疲れ切ったと嬉しい悲 鳴を上げていた。お客様の反応も機械翻訳の進歩は 素晴らしい、年々使い易くなっている印象を受けた との評価を多数載いている。

今回の翻訳フェアを総括すれば、広報期間が短かった事、会期が年末の繁忙期に重なった事、必ずしも足場の良い会場ではなかった等の反省点はあるが 総体的に見て盛況、成功をおさめたと思われる。

英日・日英機械翻訳システム

コンパクトな翻訳専用ワークステーション

DUET Qt

シャープ株式会社

シャープは、業界最小、最軽量の英語日本語双方向 機械翻訳システムの発売を開始。1988年から英日機 械翻訳システムを販売し、その開発で培った高度な 翻訳ソフトウェア技術と豊富な納入実績をもとに日 英翻訳ソフトウェアを開発しました。

<日英機械翻訳システムの特長>

1 省スペースタイプのコンパクト設計

UNIXを搭載したワークステーションでは業界 最小、最軽量を誇るA4ファイルサイズのコンパ クトボディです。

2 高度な意味解析、言語処理による適切で自然な

翻訳文の生成

意味トランスファー方式という翻訳アルゴリズム を採用しています。

3 充実した辞書による翻訳

基本語辞書は、70,000語を内蔵しています。また、 学習辞書により、訳語の優先度を変更できます。 ユーザ辞書は、ユーザ特有の言葉を新規に登録す るための辞書で、最大100,000語を登録できます。 専門用語辞書(オプション)も用意しています。



シャープ 英日・日英機械翻訳システム DUET Qt

4 1台で英文・和文の両方が入力できるOCR

英文認識・和文認識を1台で実現するOCRを接続することで、翻訳原文を効率的に入力できます。自動認識処理は、自動的に文字領域を認識できる機能で、文章の入力作業が大幅に改善されます。さらに最大30枚の翻訳原稿を一度にセットできるオートドキュメントフィーダを接続することによって、一括自動入力処理ができます。

5 多彩な翻訳支援機能による適切な翻訳

取扱説明書に多い主語のない文の訳出(命令文か 受動態)など、用途や目的に合わせて翻訳文の文 体を指定することができます。

*文節の"係り受け"を変えて翻訳する

「係り受け変更機能 | の例

「私が検討した項目を選んだ」という文章の場合

- →「私」の述語として「検討した」に係る場合
- The item which I examined was selected.
- →「私」の述語として「選んだ」に係る場合

I selected the examined item.

*言葉の切れ目を指定して翻訳する

「区切り指定機能」の例

「電動機動力」の場合

→「電動」と「機動力」に区切る場合

Electricaly powered mobility

→「電動機」と「動力」に区切る場合

The electric moter power

などの翻訳支援機能により、適切な翻訳が可能に なります。

- 6 コンパクトなボディながら高速の翻訳を実現 1時間にこなせる翻訳量は、約5,000語 (A 4 サ イズ用紙約7枚分)
- 7 パソコンやワープロなど多彩な機器と連携が 可能

<仕 様>

	ケー			
		CPU		3 2 ビットCPU (MC 6 8 0 3 0)
Λ		メモリ		1 6 MB
7	1.21	ハードディスク		2 0 0 MB
ř l	本体	フロッピーディ	スク	内蔵3.5インチ
ドウ		表示装置		液晶ディスプレイ
エ		サイズ/重量		幅320 mm×奥行き273.5 mm×高さ62 mm/3.7 kg
ア	マウス	3 ボタンメカニ	カルマ	アウス
į	基本ソフト	OA/UX4.0(UNIX	Syste	emVリリース4.0ベース)
		基本語辞書	約7(0.000語
且	辞書	ユーザ辞書	最大	100.000部/辞書(強する内容やハードディスクの残量によって制限されます。)
日英翻訳		専門用語辞書 (オプション)	情報如	四理、電子工学、機械工学、経済
ソ		翻訳方式	意味	トランスファー方式
フ	翻訳	翻訳速度	約5.	000語/時間、A4用紙(640語)約7ページ
1		翻訳処理モード	対話	型翻訳/バッチ翻訳
				省略文(受動態/命令文)など
	支援機能	区切り指定、係	り受け	一変更、別解釈、未登録語検索、検索置換など
	辞書	基本語辞書	約79	9.000語
		ユーザ辞書	最大	4 0.0 0 日語/辞書(登録する内容やハードディスクの残量によって制限されます。)
-1-1		学習辞書	利用可	可能
英日翻訳		専門用語辞書	情報如	四理、電子工学、機械工学、経済、化学工学、航空
翻		(オプション)	科学-	一般、医学一般、バイオ、自動車、コンピュータ
訳		翻訳方式	意味	トランスファー方式
ソ	翻訳	翻訳速度		2.000語/時間、A4用紙(400語)約30ページ
フ	W/1 G/(型翻訳/バッチ翻訳
۲		翻訳モード	訳文法	形態(ですます調/である調)など
	支援機能			羽訳指定、フレーズ指定、無翻訳指定、品詞指定、
	A JACTISCHE	文型指定、別解	釈、ス	スペルチェック、検索置換など

<英文・和文OCR>(オプション)

	英 文	和文
認識文字種類	英字/数字/記号(28種類)計90種類	ひらがな/カタカナ/英字/数字 記号/漢字(JIS第1水準+164文字)
認識フォント	オムニフォント	
文字フォント	6ポイント~60ポイント(2m	$m \sim 2 \text{ O mm}$)
認識速度	60文字/秒	50文字/秒
認識率	99.8% (当社測定値)	9 9.0 % (当社測定値)
認識処理モード	対話型認識/自動認識	
支援機能	自動領域指定機能、英字フォント	登録機能など

(問合せ先 シャープシステムプロダクト(株) 電話03-3267-5753 (ダイヤルイン)

英文作成支援システム

ReadyPen V02

富十通株式会社

●ReadyPenとは

ReadyPenは、ビジネスレターや契約書など海外と の取り引きに欠かせない英文書の作成を支援するシ ステムです。

既存の文書や例文を上手に使って、英文ビジネスレターをサッと仕上げるのは、国際派ビジネスマンの必須テクニック。このような 「英借文」 式英作文をシステム化しました。英文の入力・編集・印刷などの基本機能に加えて、「英借文」 データベースの参照機能、辞書引き機能、英文チェック機能などを備え、トータルな英文作成環境を提供します。

●ReadyPenの特長

[1] 【英借文】のシステム化

既存の文書や例文をちょっとだけ直して利用する 「英借文」式英作文。ReadyPenはそのために必要な 以下の支援機能を備えています。

□英文作成機能□

基本的な英文入力・編集・印刷機能を備えた英 文エディタです。さらに、英文ビジネスレターを 作成するのに便利な自動レイアウト機能や住所録 機能を利用することもできます。

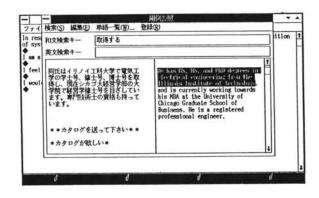
□定型文書参照機能□

ビジネスレターなどの定型的な英文書を、文例 をもとに作成します。文書の目的に応じて定型文 書データベース (定型文例集) から文例を選び出 し、ほんの一部をかえるだけで、立派な英文書が 書けます。



□用例参照機能□

和英対訳の用例データベース (用例集) から例 文を検索し、適切な英語表現を調べます。あなた の言いたいことにぴったりの英文が見つかったら 作成中の文書にコピーして活用しましょう。



□辞書引き機能□

約6万語の見出し語を持つ和英辞書を標準装備。 英文を書きながら、随時辞書引きが行えます。

[2] 強力な英文チェッカ

オプションの英文チェッカがあれば、英論文など 一般の英文作成にも大活躍。スペルはもちろん文法 や文体もしっかりチェックしてケアレスミスを防ぎ ます。また、読みやすさの指標を参考に、読み手の レベルに合わせた英文を作成することができます。

[3]優れた操作性

日本語 Microsoft®Windows™ V3.0 の優れたグラフィカル・ユーザ・インタフェースを利用して使いやすさを追求。また、パソコン通信ソフト等、他のアプリケーションとの連携も容易に実現できます。

[4] 英文書の再利用

あなたの引き出しの中には、これまで使った英文がたくさん眠っているはず。これを有効に使わない手はありません。ReadyPenの定型文書データベースや用例データベースにどんどん登録して、英文書の再利用にお役立てください。

●オプションデータベース

□ReadyPen/ビジネスレター□ (下図②)

ビジネスレター分野の定型文書データベースと用 例データベースです。出典は、「組立式英文ビジネ スレター辞典」(松本安弘氏・アイリン氏共著)で す。

□ReadyPen/技術文書□ (下図3)

技術分野の用例データベースです。出典は、「工業英語ハンドブック」シリーズと「和英てにをは発想辞典」(岡地栄氏著)です。

□ReadyPen/専門用語辞書□ (下図5~10)

分野別和英辞書です。必要に応じて、専門用語辞 書と基本辞書を組み合わせて使うことができます。

●動作環境

□ハードウェア□

・本体:富士通… FMR、FMTOWNS シリーズなど NEC…PC9801シリーズなど

(日本語 Microsoft® Windows ™ V3.0対応機種)

- メモリ:4MB以上
- ハードディスク:下図参照

□必須ソフトウェア□

- · 日本語 MS-DOS® V3.1
- · 日本語 Microsoft® Windows ™ V3.0

●ReadyPen基本セット

ReadyPen V02とReadyPen/英文チェッカ、Ready Pen/ビジネスレターをまとめたセット商品です。 一枚の 3.5インチ光磁気ディスクで提供します。

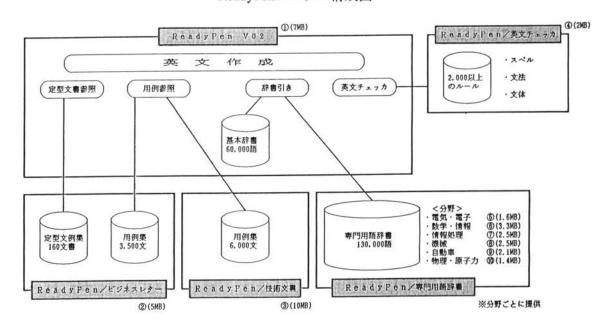
□特長□

- ・約 100MBの書き込み可能な未使用領域があります。ここは、ご自由にお使いになれます。Ready Pen のユーザーデータベースの他、日頃よくご 使用になるプログラムを入れておけば、ご自分 のシステムをいつもポケットに入れて持ち運ぶことができます。
- ・大切なシステムデータベースは、ROM化して ありますので、間違って消去してしまうことは ありません。
- ・一枚の媒体にまとめてありますので、面倒なインストール作業は不要です。
- ※ReadyPen基本セットをご利用いただく場合は、 3.5インチ光磁気ディスクドライブが必要です。
- ※ReadyPen基本セットの対応機種については、直接下記お問い合わせ先までおたずねください。

●お問い合わせ先

富士通株式会社ReadyPen係 TEL03-3730-3235 (直)

ReadyPenシステム構成図



英日翻訳支援ソフトウェア

LogoVista E to J

ロゴヴィスタ株式会社

【LogoVista E to Jの概要】

国際化の進展に伴い、増大し続ける英文情報を迅速かつより正確に翻訳・理解することは、ビジネスの最前線で活躍される方にとって必要不可欠となってきています。

LogoVista E to Jは、米国ハーバード大学の著名な 言語学者久野教授の言語理論に基づき開発された英 日翻訳支援ソフトウエアです。

「高い英文解析精度」、「的確な意味処理」「こなれた和文生成」を特徴とし、翻訳結果は<u>使える日本語と高く評価されています。</u>

翻訳精度の鍵を握る辞書は、精緻な文法規則に対応した「基本辞書」のほか、ビジネス・科学技術の専門用語を収録した「専門辞書」、ユーザ独自に構築できる「ユーザ辞書」で構成されます。

本システムはまた特定のハードウエアに限定されない「オープンシステム」として設計されていますので、既にオフィスに導入されている数多くのワークステーションやパソコンでご利用いただけます。

【システムの主要な言語的特徴】

翻訳結果の品質が翻訳支援ソフトとして、最も重要な点であることは申し上げるまでもないことですが、標準的な評価基準がないことも事実です。

ここでは実際の翻訳事例を使ってシステムの言語 的特徴を紹介します。

a) 時制・アスペクトに対する配慮

He is coming at ten. 彼は10時に来る。

He was coming at ten. 彼は10時に来ることになっていた。

- b) 「は」と「が」の使い分け
 - I know that John is a student. 私はジョンが学生であることを知っている。

c) 助動詞の意味

You may not go. あなた (達) は行ってはならない。

- I may not go. 私は行かないかもしれない。
- d) 否定形の取り扱い

Nobody came. 誰も来なかった。

- I bought few books. 私はほとんど本を買わなかった。
- e) 比較構文

John put more bananas in boxes than Mary did. ジョンは箱にメリーがそうしたより更にもっと多くのバナナを入れた。

f) 引用構文

She is sick, said her mother this morning. 彼女は病気である、と今朝彼女の母親が言った。

- g) 時間·日·年
- I get up at six in the morning. 私は朝6時に起床する。
- I was born in 1961. 私は1961年に生まれた。
- h) 意味分析による訳語の使い分け

He wears sunglasses when he goes out.

彼は彼が外に出る時サングラスを<u>かける</u>。

He was wearing cologne. 彼はオーデコロンを付けていた。

i)分裂文構文

It is always with Jane that I go to the movies.

私が映画を見に行くのは常にジェインと一緒

である。

j) 未登録語の処理

Xxx yyyed zzzly.
xxx がzzzly にyyyed した。

- k) さまざまな構文への対応
- · there 構文

There ought not to be any worry. (どの・少しの) 心配もあるべきでない。

・happy 〜that〜とas〜as〜構文の組み合せ He is as happy as Mary is that you came here.

彼はメリーがそうであると (比べて) 同じ位 あなた (達) がここに来た事が嬉しい。

· 先行詞欠落関係代名詞構文

This is what I am interested in doing.

これは私がすることに興味を持っている事物 である。

· 所有形関係代名詞構文

John will climb the mountain the top of which he knows is covered with snow.

ジョンは彼がその頂上が雪で覆われていることを知っている山を登るであろう。

- believe to be 構文
 He is believed to be dead.
 彼は死んでいると信じられる。
- 複雑構文

How easy did John become for you to talk with?

ジョンはあなた (達) にとって話しをすること がどれぐらい容易になったか?

· 分詞構文

Getting angry, John hit Bill. 腹を立てて、ジョンはビルをぶった。

感障構文

What a big boy he is! 彼は何と大きい少年であることか!

・アメリカ英語の特殊な構文

All you have to do is apologize

to her.

あなた (達) がしなければならないすべては

彼女に謝罪することである。

【システムの主な機能】

ワークステーション版、パソコン版とも同一機能 を実現しました。プラットフォームは翻訳の需要量 に応じて選択してください。

a) システムの操作性

b) 翻訳処理の基本的流れ

システムの基本操作はマウスによるプルダウン メニューの選択で簡単に操作できます。従って、 コンピュータの専門知識が無くても利用可能です。

- ・データ入力
 - キーボードから直接入力できますが、ビジネスでご利用の方にはOCRで読み込んだテキストファイル入力をお勧めいたします。
- ・翻訳環境など設定 次のような環境設定ができます。 別解釈の最大数 (~20まで可能) 和文の文体 (~だ調、~である調、~です調) 固有名詞表示形式 冠詞の訳

文書タイプ (会話体、普通体) テキストレイアウト (対訳、フォーマット済) 専門辞書 (同時に複数可能)

翻訳実行

入力されたテキストの翻訳を行います。現在翻訳を行っている文の進捗状況とテキスト全体に 占める進捗状況が画面に表示されます。各文の 翻訳が終れば逐次その結果が表示されます。

後編集

翻訳結果をみながら別解釈機能を使ってより正確な和文に仕上げます。別解釈機能では、文全体を対象に別解釈を求めることができるほか、語句を指定 (ハイライト) して部分的に別解釈を求めることもできます。このようなシステム機能を使った後編集結果は統計的に学習されますので、使いこなすほどカスタマイズされていきます。また、直接自由に修正することも可能です。

・翻訳結果をさらに良くするために、ユーザ辞書 の登録(名詞、動詞、形容詞、副詞の各語句) も簡単な操作で可能です。

(問合せ先 電話03-5690-8531)

日英機械翻訳システム

"ユーザーとの対話によって成長する"

ARGO J/E

株式会社 C S K

CSK(会長兼社長 大川功、東京都新宿区西新宿 2-6-1新宿住友ビル16階)は、過去6年間にわたり社内運用(*1)と並行して開発を進めてきた日英機械翻訳システム「ARGO J/E」をパッケージ化し、12月11日より発売を開始した。

「ARGO J/E」は、CSK独自の翻訳方式である 係り受け直接検出方式(*2)を採用しており、これまで困難とされていた、語順の自由度の高い日本 語の解析が可能になっている。

今回発売したシステムは、開発と実運用を経験したスタッフがユーザー固有の用語や言い回しの辞書 (=ユーザー辞書)を構築する、ユーザー辞書作成受託がパッケージ化されており、「購入したが十分に使いこなせない」というこれまでの機械翻訳ユーザーの認識を払拭する強力なサポート体制を備えている。

また、「ARGO J/E」は「ユーザーニーズに合わせて、使うほどに成長するシステムである」という基本コンセプトに基づいて設計されており、マルチウインドウを採用した高いGUIグラフィカル・ユーザー・インターフェース)とユーザー辞書登録ツールによって、ユーザー自身が自由にユーザー辞書に登録をすることができる、ユーザー開放型のシステムになっている。「ARGO J/E」のユーザー辞書登録機能は、名詞や動詞などの単語の他に、決まった言い回し(イデオム)や語と語の意味的つながりなどの登録が可能である。

オプションとして各分野の専門辞書(情報処理用語、地名、人名など現在7分野)が用意されている。「ARGO J/E」は翻訳の際にこれらの専門辞書やユーザー辞書を同時に複数使用することができるため、分野の異なる文章も一度に翻訳することができる。

価格は、1セットにつき基本ソフトウエア200万円、ユーザー辞書作成受託360万円の計560万円。オプションとして各分野の専門辞書がある。

今後は、日本サン・マイクロシステムズ(株)の製品の販売チャネルを通して5年間で2000セットの販売を目指す。

本システムの動作環境は以下のとおり:

ハードウエア : SUN SPARC station 2

メインメモリ:32MB スワップ領域:100MB

ハードディスク容量:300MB以上

ソフトウエア環境 : SUN OS 4.1.1 (OS)

Open Windows 2.0.1以上

(ウインドウ環境)

JLE1.1以上(日本語環境)

以上

(*1) 社内運用:

日経テレコンの金融・証券記事の英訳 サービス

(*2)係り受け直接検出方式:

文中の単語間の関係に注目して、両者 の修飾/非修飾関係をしらべながら文 の意味を表す表現 解析木)を作ってい く解析方法。修飾/非修飾の決定には、 品詞、意味素、格、属性などが利用さ れる。

■本件に関するお問い合わせ先:

CSK社長室広報グループ/小林・宮地(03-3505-5504)

■製品に関するお問い合わせ先:

CSK技術開発本部 MT開発部/永野(03-3342-3047)

AAMTジャーナル 1993.2 No.2

台;双方向翻訳 今;一方向翻訳 /;相互間翻訳

現行製品一覧

惠壱)

信田

(作成者:シャープ(株)情報技術開発センター

[MT SYSTEM]

		メーカ名称 及び 開発研究機関名称	システム名	翻訳対象言語
	-:	Advanced Telecommunications Research	[Interpreting Telephony Research]	
20175	5.	Catena-Resource Laboratories Inc.	[STAR] [The Translator]	EウJ
38116V	3.	Center of the International Cooperation for Computerization	[MTS for ASIAN LANGUAGES]	J/Chi/タイ/インドネシア/マレー(相互間)
120	4.	CSK	[ARGO-J/E]	J今E
(8.7)	5.	Fujitsu, Ltd.	[ATLAS]	E⇔J
.55	9	GLOBALINK(Iris International, Inc.)	[GTS]	EGG EGF EGSp EGRu
	7.	Hitachi, Ltd.	[HICATS]	
20000	<u>«</u>	IBM Japan, Ltd.	[SHALT/JETS]	E⇔J
8008	9.	Japan Electronic Dictionary Research Institute, Ltd.	「電子化辞書』	
.—	0.	LogoVista	[LOGOVISTA]	EウJ
-	Ξ.	Mitsubishi Electric Corporation	[MELTRAN-J/E]	J·今E
-	15.	Matsushita Electric Industrial Co., Ltd.		JOE
	13.	NEC Corp.	[PIVOT]	Е⇔Ј
-	14.	NOVA Corp.	[PC-Transer]	ЕӨЈ
-	15.	NTT Communications and Information Processing Laboratories	[ALT-J/E]	J·今E
	16.	Oki Electric Industry Co., Ltd.	[PENSEE]	Е⇔Ј
-	17.	Osaka Gas Information System Research institutte Co., Ltd.	[PENSEE]	E⇔J
,—,	18.	RICOH Co., Ltd.	[RMT]	EウJ
	19.	SANYO Electric Co., Ltd.	[HEAVEN]	E⇔J
6.4	20.	SHARP CORPORATION	[DUET-Qt]	E⇔J
64	21.	Toshiba Corporation	[ASTRANSAC]	ЕфЈ
6.4	22.	NIFTY-Serve		Ecl Ecc Ecr Ecsp
	23.	PC-VAN		ЕфЈ
MT?	24.	EYE-NET(Fujimick)		Е⇔Ј

E.英語 J.日本語 Chi.中国 F.フランス G. ドイツ Sp.スペイン Sw.スウェーデン It.イタリア Du.オラング Po.ポルトガル Ko.韓国 Ru,ロシア語

AAMTジャーナル 1993.2 No.2

システム評価研究会活動報告

電子技術総合研究所 横山 晶一

システム評価研究会は、田中 穂積委員長(東工大)のもとで、機械翻訳の評価を行なう目的で設立された。当初、ユーザーの立場から機械翻訳システム全般の評価を行なう環境評価ワーキンググループ < WG > (主査: 牧野 武則副委員長(東邦大))と、自然言語処理の観点から機械翻訳システムを評価してみようとする言語評価WG(主査: 横山)とに分かれていたが、前者が研究会として独立したため、現在後者の観点からの活動を行っている。委員会のメンバーは、当初17人、環境評価WGが独立した時点で言語評価WGに属する8人が残り、新しいメンバーを加えて現在16人である。構成は、メーカーとユーザーが大体半々といったところである。

発足当初は、まず有識者の意見を聞くということで、池原 悟氏(NTT)、水谷 静夫氏(IBS、東女大名誉教授)、坂本 義行氏(電総研)、石綿 敏雄氏(日女大)からそれぞれご意見を伺った。その結果については、昨年6月のセミナーの通りである。

92年度は、自然言語処理の観点から機械翻訳システムの評価をすべく活動を開始した。 まずユーザー、メーカーの双方から、機械翻訳をする上で、翻訳が困難である例文を提供してもらった。今年度は、まず例文を日本語文とした。その理由は、入力文について種々の分析を行う場合には、内省のできる日本語の方が好都合であると判断したためである。次に、これらの例文を分類した。分類は、日本電子工業振興協会(電子協)が昭和60年に行った機

- ・バブルの崩壊もあり、不景気が続いている。
- · A bubble decays, too, and a depression continues
- \cdot There is collapse of \mathcal{NIN} , too and the depression continues.
- There is a collapse of a bubble too, and the depression is successively.
- · There is also a collapse of a bubble, and a depression is continuing.
- · The bubble collapses and the recession continues.
- #Business has been dull since the "bubble" economy ruptured.

図1 翻訳困難例文とその翻訳例 (#は人間の訳) 械翻訳例文調査の分類を参考にして、新たな分類を付け加えたものである。これらの例文を原則として辞書登録や前処理を行わずに、日英の機械翻訳システムにかけて英語を出力した。例文の1例と、その機械翻訳による訳例を図1に示す。

翻訳困難例文を無理に翻訳しようとするのであるから当然いろいろな場所でいろいろな「誤訳」が生じる。これらを分析して、自然言語処理の未熟さに起因する問題点、機械翻訳の方式に起因する問題点、言語間の構文、意味の異なりに起因する問題点などを明らかにしようとしている。

また、提供された例文に対する言い換えを(人間の手によって)作成し、その言い換えに対する機械翻訳出力を得ることによって、上記の問題点を明らかにするという作業も進めている。図2に言い換えの例を示す。

さらに、池原氏が機械翻訳システムの評価のため に作成した評価用例文、分類との比較を行っている。

現在1~2か月に1回会合を開いて上記の問題点を検討しているが、資料の整理に思いのほか時間がかかり、今年度は問題点の洗い出しにとどまりそうである。

今後の活動方針としては、これら問題点の中から 自然言語処理の克服すべき点と、機械翻訳のそれと をまとめ、翻訳例文の基本セットの作成、翻訳困難 例文の言い換えの指針などを目指している。現在ま でに提供された機械翻訳例は、貴重な資料となるの で、今後何らかの方法で公開することを考えている。

- ・バブルの崩壊も原因となって、不景気が続いている。
- ・バブルの崩壊などの原因で、不景気が続いている。
- ・バブルの崩壊にも起因して、不景気が続いている。
- ・バブルの崩壊も関係して、不景気が続いている。
- バブルの崩壊によって、不景気が続いている。
- ・バブルの崩壊の例のように、不景気が続いている。

図2 翻訳困難例文の言い換え例

沖電気工業株式会社

総合システム研究所 羽下雄之輔

当社には「知力工学研究会」なるものがあります。 知力工学 (Wits Engineering) と は「人間の創造・生産活動に必要な能力を、知力と 捉え、その力の特性を工学的に解明し、それによっ て知力の強化策を講じることを可能とする学問」と 定義され、知力発生源である知力システムは「その 力を発揮するとき、仮に知流(概念の動き)が流れ ているとすると、その媒体は言葉である。従って知 力システムの基本媒体である言葉について理解して おく必要があろう。即ち人間は、相互のコミュニケ ーションのためのツールとして言葉を作ったが、そ れは100万年の間に進化し、もはやツールの概念 を超えて、知的活動の媒体 (メディア) として人間 の頭脳の中に取り込まれてしまった。従って言葉の 問題、即ち言語現象を、工学的に理解しておくこと は知力強化の立場からは重要な出発点となるのであ る。同時に言語能力強化が知力強化の基礎にもなる のである。」(以上、大須賀監修、山本編「知力工 学のすすめ」オーム社、1992より引用)

言語現象の工学的研究として、機械翻訳が最も適切なものであると当社は早くから認識して研究を行ってきました。これまでの研究成果として日英双方向の機械翻訳システムPENSEEの開発を通して、日本語と英語の特徴が明らかになってきました。日本語の語彙数は、英語に比較して多く、英語の語彙は抽象的な語義を持つ単語が多いこと、構文的には日本語は自由度があり、英語は制約が厳しいことなどです。機械翻訳に必要な解析規則はどちらも同程度(約1500ルール)であることが分かりました。現在の研究課題としては、文法規則の自動学習や語彙知識の自動獲得などがあります。

自然言語処理関連の当社の研究は、研究開発本部の関西総合研究所および総合システム研究所で行なっています。また研究開発本部商品化推進部では研究成果を事業につなげるための開発研究を行なっています。

PENSEEの研究開発は、関西にある関西総合研究所で行っており、大阪ガス(株)殿、(株)オージス総研殿との共同研究であります。PENSEEの研究は既に10年になり、最近は訳質の向上だけでなく、使用環境の利便性の向上を進めています。

たとえば、FAX入力-OCR認識-PENSEE翻訳の組合せ、音声認識と合成技術を利用した音声入出力付き機械翻訳、ネットワークを介した機械翻訳サービス、富士ゼロックス(株)殿の文書作成システムGlobalView上にPENSEEを搭載して文書作成から翻訳、印刷までを統合する機械翻訳付きのDTP等々、多様な翻訳サービス機能を研究・開発しています。

東京にある総合システム研究所では、自然言語の基礎的研究を中心に、第5世代コンピュータプロジェクト、EDR電子化辞書プロジェクト、CICC多言語機械翻訳プロジェクト、フレンド21プロジェクト等自然言語処理、知識処理の国家プロジェクトに参画することにより、将来の基礎固めをするとともに、自然言語処理の事業化をにらみながら、オープンシステムの発想で新しい応用分野を模索してます。

従来、自然言語処理システムはワードプロセッサや機械翻訳に代表されるように言語システムとして閉じた系で開発されてきました。一方で、情報システムを考えますとその知能化には知識や言葉といった知的なメディアが必須であることは明らかになってきています。最近の情報システムの動向からしますと、自然言語処理が飛躍するには言語単独の世界から広くマルチメディア情報の世界と協調することが必要だという状況が到来するものと予想しています。

自然言語処理のアプローチとして、言語情報および 文書情報の組織化・構造化に注力しています。辞書は 語彙構造を規定し、文法は文構造を規定するものとみ ることができます。文章や会話の構造化としては談話 理解の研究があります。具体的には、構文構造の研究 として汎言語的枠組みで研究を続けており、文脈構造 の研究として、新聞社説の構造化で成果を得ました。 文書情報の構造化としては、文書レイアウトの自動認 識やハイパーメディアでの文書構造の表現法を研究し ています。今後の研究方向としては自然言語理解のた めにコンピュータビジョンや心理学、認知工学といっ た 学際領域 の研究も 必要であると考えています。 他 方、企業としての事業化研究として、いままで試作し てきたいくつかの自然言語処理システムのプロトタイ プを、オープンシステム環境下で再構築していくこと を目指しています。

ユーザー事例紹介

株式会社 十印 言語処理部

1.はじめに

弊社は'64年に設立され、創業当初より技術翻訳専門会社として知られていると思いますが、現在ではマニュアル制作会社の色彩の方がより濃くなっています。

さて、そういう会社がどうして機械翻訳などという未だに賛否相半ばするソフトウェアを利用して、 日本語マニュアル制作を開始するに至ったのか、お 話したいと思います。

2. これまでの経緯

さきほど、翻訳からマニュアルへと事業の中心が移ってきたと書きましたが、実は、もう一つ弊社の事業の中で、'83年以来傍流として流れ続けているのが、機械翻訳の開発業務なのです。これは、勿論弊社独自の開発ではなくて、開発メーカさんから仕事を請負って、主に翻訳辞書と翻訳規則などの自然言語データを開発する業務です。

そうこうする内に、'88年に開発業務側の発案で弊社の三つのシーズ、即ち、技術翻訳・マニュアル制作・機械翻訳を組み合わせて、まったく新しい方式の翻訳マニュアル制作ができないかということでとりあえずプロジェクトチームを発足させました。しかし、それぞれの事業から要員が結集したわけではなく、当面は機械翻訳の開発側だけでプロジェクトを推進することになりました。残りの二つの事業部から人的・技術的支援を受けたのは、ずっと後のことになります。それは、機械翻訳そのものに対する当時の社内の評価は、「夢の商品のままでいいのではないか」だったからでしょう。

ともあれ、開発側としては新しい方式によるマニュアル制作にこだわりをもっていましたので、機械 翻訳を夢から現実のものにするために、諸問題を検

討し、

- (1) 英日翻訳から取り組む
- (2) 磁気化された媒体のみ扱う
- (3) ユーザ辞書とユーザ翻訳規則を構築する
- (4) マーケティングを強化する
- (5) 生産工程を標準化する

を一つ一つ解決していくことにしました。少し補足 しますが、

(1)は、マーケットを考えれば日英というのも捨て 難いのですが、(3)との絡みでターゲットの言語が英 語ですと、どうしてもネーティブのスタッフが必要 になるでしょうし、開発側が目標にする翻訳品質の 設定そのものに、英日より時間を要するだろうとい う理由からです。

(2)については、 紙の媒体も検討しましたが、 翻 訳システムへ取り込む上で、紙のオリジナル情報と OCRや手入力などで磁気化された情報を照合するという余計な工程・コストが発生するため、磁気媒体 のみに限りました。

(3)は既存商用システムを、ターゲット分野用にチューニングするな作業で、これは日立製作所殿の全面的なバックアップがあってこそ実現できたことです。このチューニングを通じて、当面の機械翻訳が目標とする翻訳品質が定まったといえると思います。

(4)については、コンピュータという機械を利用して翻訳するわけですから、対象となるマニュアルは少量・多品種ではなく、大量・少品種であることが条件となります。そういう条件下ではじめて機械の強みが発揮されるわけですので、その線でマーケティングを行いました。

(5)は当たり前のことですが、試行錯誤の結果、大まかに言って、ユーザ辞書の登録、磁気媒体の前処理、機械翻訳、リライト(ポストエディト)、後処理、版下という流れになっています。

そして、マーケティングの甲斐もあり、また世の中の機械翻訳に対する風向きも少しずつ変化してきたこともあって、'91年早々日本語マニュアル制作を初めて受注することができました。

3.現在の到達点

ビジネスを開始して、二年経ったわけですが、現 在、弊社が立っている地平について述べてみたいと 思います。

弊社のコンセプトを一言で表すと、「機械翻訳を 利用したマニュアル生産」ということになるかと思 います。ここで「制作」でなく「生産」という言葉 にしたのは、二つの理由があります。

一つは商品価値の中心部分を占めているのが、翻訳であり、その翻訳の前段部分を機械が担っているわけです。そこで機械がもつ「生産」というイメージに弊社はこだわりたいのです。

もう一つは人間の関与する作業を極小化するというテーマで取り組んで来ていますので、生産体制全体を工場的なイメージで捉えてていただきたいという願いからです。

さて、生産体制について少しお話します。弊社には日立のワークステーション3050と2050があり、このマシーンで機械翻訳システムHICATS/EJを稼動させ、翻訳生産とユーザ辞書管理・作成を行っています。HICATSに取り込む前の前処理、リライト後の後処理にはPCを利用しています。さらに版下を作成するためにHP、SUNのWSやIBMPCLBPがあります。これらは全て社内ネットワークで結ばれており、一連の工程に従って流れ作業的にデータが加工されていきます。リライタにはパソコン通信を利用してHICATSが訳した翻訳データを送り、リライトされたデータを送り返してもらっています。従って社内のスタッフにはリライタはひとりもおりません。

現在の到達点を版下中心に書いてみますと、弊社は、テキストのみの磁気媒体は原則として取り扱いません。これは、翻訳のみでは付加価値が低いということもありますが、弊社の生産方式が元々、ソースの磁気媒体に存在するマニュアル上の情報を全て保存して日本語マニュアルに再現するという考え方

からきているからです。もう少し詳しく言いますと、ソース側のデータにはテキストは勿論のこと、レイアウトやイメージ、テーブル、フォント、インデックスの情報が存在していて、そういう情報も丸ごとHICATSの世界に取り込み、日本語側で再現すれば、版下作成の生産性が非常に高まるからです。こういう版下作成方式に拘ったのは、やはり、「生産」というコンセプトがあったからです。弊社が処理可能な版下(組版ソフト)は、今のところ、マークアップ言語、エムエスワード、フレームメーカ、インターリーフの四つです。この四つに絞っているのは、マーケティングの結果でもありますが、いちばん大きな理由はこういうソフトで作成された磁気媒体が現在の生産方式に完璧にのるからです。

4.今後の課題

課題はたくさんあります。取り上げたらきりがあ りませんので、一つだけ書きます。それはリライタ の育成です。いくら工場生産的な体制を作り上げて も、リライトするのは生身のリライタですから、彼 らとの協力関係がなければ、このビジネスは成立し ません。リライタに渡すデータはHICATSが訳出した 部分と訳出不要のシンボル化された部分が混在して いますが、テキスト部分以外は、やはり慣れないと かなりのノイズになっていると聞いています。不要 情報を無視しながら機械翻訳らしい翻訳文をリライ トするのはやはり慣れるまでは、ある程度の時間を 要します。しかし、一方で翻訳品質の目標値を着実 に高めていこうとする努力と翻訳システムが経年的 に向上する見込みがあることを考え合わせると、リ ライタビリティは確実に上がっていくと考えていま す。

またノイズについては、目下、前処理ツールの改 良を続けており、低ノイズ化を目指しています。こ ういう問題をクリアすることによってリタイヤの層 を厚く広げて行きたいと考えています。

(株式会社 十印 言語処理部長 鳴海武史)

第6回

文生成技術

今回は、文生成の技術について解説します。機械翻訳における文生成とは文字どおり目標言語の訳文を生成する処理のことを指します。まず、機械翻訳の文生成に限定しないで、より広い観点から自然言語を生成するための基礎的な技術について説明し、その後で、機械翻訳に固有の文生成の話題について説明します。

Q:自然言語生成って何ですか?

A:自然言語生成は、「入力として与えられたゴールを達成するためにテクストを生成すること」と定義できるでしょう。ここでいうテクストは文章に限らず、対話における発話なども含みます。重要なことは、自然言語をあるゴールを達成するための道具としてとらえることです。

A:自分で文章を書くことを考えてみてください。 まず、何について書くかということ、つまり内容を 決めなければなりません。たとえば、日記を書く場 合を考えてみましょう。日記にはその日に起ったこ とを書くわけですが、その日に起ったことすべてを 書くわけにはいきません。当然、取捨選択が必要に なります。 たとえば、 その日のうちで一番印象深 かったことだけを書くというのもひとつのやり方で しょう。これがすなわち内容の決定です。

内容が決まったら、それをどのような構成で述べるかを決めます。たとえば、多くの推理小説では、まず事件が起り、探偵が登場し、探偵が犯人を探す、という構成になっています。しかし、刑事コロンボのように、まず、犯人の犯行を見せておくという構成も考えられます。内容的には同じでも、このように構成を変えることによって、小説全体の印象はずいぶんと変るわけです。

さて、内容が決まって、構成が決まったら、次に どのような言語表現を使ってその内容・構成を表現 するかを決めなければなりません。生成する言語の 文法に照らして正しいテクストを生成することはも ちろんですが、同じことを表現するのにどのような 語を使うのか、受動態を使うのか、能動態を使うの かなど様々なことを決めなければなりません。この ためにはテクストの読み手あるいは聞き手がどのよ うな知識を持っているのか、あるいは自分との関係 はどうなのか、どのような種類のテクストを生成し ようとしているのかなど、様々な要因を考慮する必 要があります。たとえば、「怪盗ルパン」という本 には、大人向けのものもあれば小学生向けのものも あるわけです。同じ内容・構成でも、その表現はず いぶんと違っています。

以上をまとめると自然言語の生成には以下の3つの処理が必要だということがわかります。

- ・表現する内容の決定
- ・内容の構成の決定
- ・ 言語表現の決定

内容の構成を決定する上では、さらに内容を追加したり、削除したりすることもあるので、内容の決定と構成の決定の境界は必ずしも明確ではありません。自然言語生成の分野では普通、これら2つをまとめてwhat-to-sayの決定、あるいはdeep generationと呼びます。一方、言語表現の決定はhow-to-sayの決定、あるいは surface generationと呼ばれます。

ここで注意して欲しいのは、これらの処理は必ずしも独立にこの順序でおこなう必要はないという点です。解析において形態素処理、構文処理、意味処理などを融合しておこなう手法があるように、生成でもwhat-to-sayの決定とhow-to-sayの決定を融合しておこなうこともできます。自然言語生成では実際にテクストが生成されるまでに、このように様々な決定をおこなわなければなりません。これらの決定をうまくおこなって、より質の高いテクストを生成することが自然言語生成の目的です。

Q:あれ、最初の答では、自然言語生成の目的は与えられたゴールを達成することだということでしたが。

A:そうです。質が高いテクストを生成するという ことは与えられたゴールを達成するための手段です。 したがって、厳密にいえばテクストの質は現在のゴー ルをどの程度達成できるかによって測られるべきで しょう。このように考えると、読みやすくて内容が 理解しやすいからといって、それが質の高いテクス トだとはいえなくなるような状況もあります。たと えば、佐川スキャンダルの証人喚問における証人の 答弁は、一般的に云って決っしてわかりやすいもの ではありません。*しかし、生成する立場からすれば、 一見、質問に対する答えを述べているようにみせか けながら、できるだけその内容を聞き手に理解させ ないというゴールを持っているわけですから、ゴー ルの達成度という観点から考えると、あのような答 弁は質の高いテクストだということになります。し かし、多くの場合、読みやすくて内容が容易に理解 できるようなテクストは質が高いと考えてよいでし よう。

Q:自然言語生成において入力の「ゴール」が重要だということはわかるのですが、これは具体的にどういうものなのか、いまひとつイメージがわきません。

A:その点がまさに自然言語生成の研究を困難にしている原因です。自然言語解析では入力は文字列と昔から決まっていて、それに疑問を抱く人は誰もいません。これに対して自然言語生成では何が入力になるのかが必ずしも明確ではありません。冒頭で述べたように「なんらかのゴール」が自然言語生成処理への入力であるという点では、現在、多くの研究者が合意しています。しかし、このゴールが具体的に何なのか、どのように表現されるのかについては何の合意もありません。たとえば、このゴールを

「与えられた構文木を表層の文に変換すること」とすれば、生成の処理は構文木の葉をたどるだけの非常に単純な処理になってしまいます。 一方、佐川スキャンダルの例であげたようなゴールを達成するためには、様々な要因を非常に綿密に考慮する必要があり現在の技術では困難でしょう。このように

「ゴール」をどのようにとらえるかによって自然言 語生成の処理の難しさは大きく変ってきます。

Q:これまでに開発された自然言語生成システムで はどのようなものを入力としていたのですか? A:たとえば、意味構造を入力として与えられて、その意味構造を表すような言語表現を生成するシステムがあります。このシリーズでは、これまでに、形態素解析、構文解析、意味解析の解説がありました。これらの処理によって文から意味構造を抽出できることがわかったと思います。これらのシステムでは逆に意味構造から文を作りだすわけです。このためには、語を選択したり、構文構造を決めたりする処理が必要になりますから、構文木を与えられて文を生成するよりは処理が複雑になります。意味解析までおこなう機械翻訳システムの生成部では、やはり意味構造が入力となりますから、同じような処理をすることになるでしょう。

このほかに、データベースへの問い合わせシステムの一部として作られたもので、質問に関連するデータベースの内容から文章を組み立てるシステムや、株価の数値データを入力として、株式の変動に関するレポートを生成するシステムなどもあります。これらのシステムの「ゴール」は、それぞれ「ユーザの必要な情報を与えること」、「株式の変動をユーザに簡潔に報告すること」だといえるでしょう。

Q:機械翻訳の訳文生成におけるwhat-to-say、how-to-say の決定では、具体的にどのような処理をするのですか?

A:機械翻訳の場合、原文の解析によって原文の表す意味内容が抽出されているはずですから、基本的にはこの意味内容が訳文の内容となるはずです。また、現在の機械翻訳システムのほとんどすべてが1文を単位として翻訳しますから、内容の構成に関しては何もする必要がないといえるでしょう。したがって、現在の機械翻訳では、what-to-sayの決定に関しては特に必要な処理はありません。機械翻訳の訳文生成ではhow-to-sayの決定が中心的な処理となります。

Q:では、how-to-sayの決定ではどのような処理を するのですか?

A:このシリーズの第1回の解説で説明があったように、機械翻訳の方式には大きくわけて、トランスファ方式と中間言語方式があります。機械翻訳の訳文生成は、どちらの方式を採用するかで処理の方法が変ってきます。トランスファ方式では、原言語を

[※]たとえば、このよううな例を引くことにより、この原稿がいつごろ書かれたかを読者に知らせるというゴールを達成できます。

解析した結果は、その原言語に依存した構造を持っています。たとえば、原言語が英語で、解析結果が 句構造で表現されるならば、その句構造は英語に固 有の構造を反映しているでしょうし、句構造の中で 使われる語も英語の単語になるでしょう。意味解析 までおこなえば解析結果は表層の言語表現から多少は独立な概念的構造になるでしょうが、それでも英語に依存したものになるはずです。したがって、トランスファ方式では、原言語の解析結果から目標言語の文を生成するために、少なくとも次の2つの処理が必要となります。

- ・原言語に依存した解析結果を目標言語に依存し た構造に変換する
- ・変換した結果から目標言語の文を生成する

前半をトランスファ過程、後半を狭義の生成過程と呼びます。ここでわざわざ「狭義の」という修飾を付ける理由は後ほど説明します。トランスファ過程は、さらに、語彙変換(lexical transfer)と構造変換(structual transfer)の2つの過程からなります。語彙変換では、構文的、意味的な情報を手がかりに、原言語の語を目標言語の語に変換します。これを訳語選択と呼びます。訳語選択は訳文の質を決める上で重要な処理のひとつです。たとえば、次の例について考えてみましょう。

- (1) I took a picture.
- (2) I took an apple.
- (3) I took a bicycle.

ご存知のように英語の"take"にはいろいろな意味がありますから、これをたとえば日本語に訳すときには意味の違いを考慮した上で正しい訳語選択をおこなわなければなりません。この例の"take"はそれぞれ、(1)「撮る」、(2)「取る」、(3)「乗る」と訳すのが適当でしょう。この例では、"take"の目的語がどのような意味的性質を持つかによって適当な訳語を選択することができます。名詞句の意味的な性質を表現するためには、「意味解析技術」の章で解説があったように、意味素性やシソーラスなどを使います。しかし、意味的な性質だけでいつも正しい訳語が選択できるわけではありません。たとえば、「懸賞に

当たって自転車か冷蔵庫を選べるような状況で、自 転車を選んだ」という文脈を考えると、(3)の"take" は「選ぶ」と訳す方が適当でしょう。どちらの訳語 を選ぶかは、この文が現れた文脈を考慮しないと何 ともいえません。

次にトランスファ過程の後半、構造変換について 説明します。構造変換は解析結果としてどのような 構造を仮定するかによって 具体的 な 処理は異りま す。ここでは、構文構造を例にとって説明します。 英語とドイツ語のように同じ語族に属する言語の間 の翻訳では、原言語の構文構造とそれに対応する訳 文の構文構造がよく似ている場合が多く、構造変換 では若干の構造の調整だけですみます。しかし、英 語と日本語のように構文構造が大きく異なるような 言語の間の翻訳では、構文構造を変換する必要がで てきます。たとえば、英語では無生物が主語になる 文はごく一般的ですが、日本語では非常に不自然な 表現になってしまいます。

- (4) Her present made me happy.
- (5) 彼女のプレゼントが私を幸せにした。
- (6) 彼女がプレゼントをくれたので、私はうれしかった。

たとえば、(4) の訳としては(5) よりも(6) の方が適当でしょう。しかし、(4) と(6) の構造は非常に異なっています。(4) は1 つの単文ですが、(6) は2 つの単文が接続助詞「ので」で接続されています。トランスファ過程の構造変換ではこのような処理をおこないます。

さて、以上のようなトランスファ過程をへて、目標言語に適した構造が得られると、次にこの構造から目標言語の訳文を生成します。 さきほどはこの過程を狭義の生成過程と呼びました。 トランスファ過程の内容をみればわかるように、how-to-say の決定に関する処理を実際にはトランスファ過程でおこなっています。したがって、機械翻訳における生成過程としては、トランスファ過程も含めたほうが適当でしょう。 そこで、これまでわざわざ「狭義の」とことわったのです。狭義の生成過程でおこなう処理は、トランスファ過程の結果として、構文構造が得られるとしたら、 狭義の生成過程の基本的な役割は、各語を一次元に並べ、形態素変化を適切におこ

なうことになります。

少し長くなりましたが、トランスファ方式で、訳 文がどのように生成されるかの概要を述べました。 一方の中間言語方式では、トランスファ過程と狭義 の生成過程の区別が明確ではない点がトランスファ 方式と異なります。しかしながら、中間言語方式でも 中間言語の語彙や構造から目標言語への語彙や構造 への変換は必要です。つまり、中間言語の表現を参 照して、訳語や構文構造を決めることになります。

Q:訳語選択では文脈が大事なこともあるということですが、文脈が重要な役割をはたす例にはどのようなものがありますか?

A:人間が英語を日本語に翻訳する場合、英語の代名詞は翻訳されずに省略されるか、その代名詞の指示している名詞が明示的に繰り返して用いられることが多いといわれています。1文を翻訳の単位として扱う現在の機械翻訳では、原言語の代名詞を目標言語の対応する適当な代名詞に翻訳し、実際にそれが何を指すのかは、出力された文を読む人間の推論能力にまかせるしかありません。たとえば、以下の例を考えてみましょう。

(7-1) Taro bought a new camera.

(太郎は新しいカメラを買った。)

(7-2) He likes it.

(彼は それが気に入っている。)

(7-2)において、「彼」が太郎であり、「それ」がカメラであることは、文を単位として翻訳している限りわかりません。(7-2)では、"he"、"it"をそれぞれ、「彼」、「それ」と翻訳しても人間が読めば「彼」、「それ」が何を指しているのかわかるのでそれほど不自然な訳とはなりません。しかし、代名詞が何を指示しているかによって、訳語が大きく影響を受ける場合があります。次の例は代名詞が何を指すかが正しく解析できないと、正しい訳語選択ができない例です。

(8-1) Taro took a picture.

(太郎は写真を撮った。)

(8-2) He developed it.

(彼は それ(フィルム)を現像した。)

(9-1) Taro have a good camera.

(太郎は素晴しいカメラを持っている。)

(9-2) He developed it.

(彼が それ(カメラ)を開発した。)

(8-2) と(9-2) はまったく同じ文ですが、先行する文によって"it"の指す対象が異なっています。 いずれの場合も"it"を「それ」と訳すことに問題はありませんが、"it"が何を指すのか正しく解析できないと、"develop" の訳語選択が正しくおこなえません。 英語の"it"をドイツ語やフランス語の代名詞に翻訳する場合にも同様な問題が起ります。 英語の"it"には性に関する情報が含まれていませんが、ドイツ語やフランス語では、代名詞は性によって違いますから、"it"が男性名詞を指しているのか、女性名詞を指しているのかによって代名詞を選択しなければなりません。さらに、英語の冠詞の"a" と"the" の使いわけ、日本語の助詞「が」と「は」の使いわけなども文脈の情報を参照しなければ、正しい選択ができません。

Q:将来の機械翻訳システムには自然言語を「理解」 することが必要なわけですね。

A:もちろん、自然言語理解の技術が完全なものになれば理想ですが、そのためには長い時間と多くの研究が必要でしょう。工学的な立場に立てば、実用上許容できる程度の翻訳を出力するにはどのような情報が必要か、またそのような情報を得るためにはどのような解析をおこなえばよいのかというアプローチで研究を進めることが重要でしょう。そのためには生成された訳文を評価する手法を確立し、どのような点が訳文の質を落しているのかを分析することが必要です。これまで、機械翻訳に限らず自然言語処理の分野では生成技術に関する研究はあまり盛んではありませんでした。しかし、より質の高い機械翻訳システムを構築するためには、これから生成技術の研究がますます重要になってくるでしょう。

(東京工業大学 工学部 徳永健伸)

「機械翻訳評価ワークショップ」報告

九州工業大学情報工学部 野村 浩郷

国際機械翻訳連盟(IAMT)とアメリカ機械翻訳協会(AMTA)が主催した「機械翻訳評価ワークショップ (MT Evaluation Workshop/Showcase)」は、機械翻訳評価の現状と将来に関する発表と討論及び機械翻訳システムの提示を目的として、米国政府の科学基金(National Science Foundation)より財政援助を受け、1992年11月2日より3日までの2日間、米国サンディエゴのプリンセスホテルで開催された。

ワークショップへの参加者は、事前登録者が113名で、その内訳は米国の他からは日本6名、英国3名、カナダ3名、スイス2名、スウェーデン2名、ルクセンブルグ1名、韓国1名である。当日参加登録者もかなりあり、最終的に130名ないし140名くらいの参加者があった。両日とも朝8時半から夕方まで興味深い発表と活発な討論が行われた。各講演の概要を収めた薄い資料が配布された。ショーケースはワークショップの隣の会場で開催され、ワークショップへの参加者の他にショーケースのみの見学も有料で許されていた。

ワークショップの閉会式では、アメリカ機械翻訳協会会長のDr. Muriel Vasconcellos による開会宣言とアメリカ機械翻訳協会の現状の説明、国際機械翻訳協会のProf. Makoto Nagaoによる世界機械翻訳連盟設立の経過と目的及び日本機械翻訳協会がアジア太平洋機械翻訳協会として拡大されたことの報告、さらに1993年夏に日本で開催予定のMT SUMMITIV及び TMIの案内と国際機械翻訳連盟が発行しているNews Letterについての紹介、ヨーロッパ機械翻訳協会会長のProf.Margaret Kingによる欧州で開催した機械翻訳評価に関する二つのワークショップの報告と現

在進行中の評価実験の紹介があった。さらに、この機 械翻訳評価ワークショップのスポンサであるNational Science Foudation から Y. T. Chein 氏が祝辞を述べ ると共に、米国における機械翻訳研究の支援状況に ついて紹介した。

米国の Prof. Y. Wilks は、米国における機械翻訳評価は成功裡に進みつつあるが詳細な評価は将来に待たねばならないこと等を述べた。欧州のProf.M.Kingは、欧州における機械翻訳評価への取り組み状況と、そのための国際協力の必要性を述べた。欧州共同体委員会(CEC)の Dr. L. Rolling は、CECにおける機械翻訳評価への取り組み状況と今後の計画を説明した。

機械翻訳評価について、日本電子工業振興協会の機械翻訳評価について、日本電子工業振興協会の機械翻訳評価基準の内容の紹介、米国のMITRE社が進めている機械翻訳システムの評価への取り組みについての報告があった。また、米国のDARPAが進めている機械翻訳や文書からのデータ抽出についての評価の報告がいくつかあった。機械翻訳の新しい方向に関して、米国及び英国から技術的な講演が行われた。パネル討論でも、評価に関するいろいろな問題点、及び今後の技術開発における問題点が議論された。

機械翻訳システム等を開発している企業や研究機 関等によるそれぞれの内部評価についての紹介が行 われた。また、評価を行うためコーパスや辞書の問 題点や開発への取り組みなどの紹介もあった。日本 からはJICSTとNTTから紹介があった。

機械翻訳システムなどの展示には、米国から11社、 日本からシャープの1社が参加し、好評であった。

(1993.2.9)

機械翻訳時代の翻訳者の役割は?

元吉 宏子



原稿用紙(実はパソコン)を前に機械翻訳についてあらためて考えてみた。締切が近づくと「ああコンピュータに原稿を読み込ませるだけで翻訳が仕上がり、後はコーヒーを飲みながらのんびりチェックすれば.. なんてことにならないかな」と想像することはよくあるが、具体的な知識はほとんどない。

衛星放送で「機械翻訳」という但し書きがついた 字幕のニュースを見たことがあるが、とても楽しい 頭の体操になった。というのは、 字幕の日本語か ら、原文の英語の輪郭がはっきり読み取れ、自分も 多分はこう訳出してから、前後を考えてもっと適切 な単語に変えるとか、文章の順序を動かすかなとい うプロセスが次々と頭に浮かんできたからである。 でも、英語を日本語におきかえるという作業に日頃 なやまされている翻訳者としては、行間ににじみで ている翻訳機の苦労にすくなからず 親近感 をおぼ え、とてもひとごとと思えない共感を感じた。

そんなことを考えながら、学生時代に読んだときは文の流れがつながらず、途中で放り出し「ジェーン・エア」の日本語訳をもう一度読んでみた。そこには、あの機械翻訳と同じようにあたかも英文を読んでいる気分になるほど「英語らしい(?)」日本語の文章があふれていた。でも翻訳者としての自分も知らず知らずに同じ間違いをおかし、お読みになる方々に迷惑をかけているんだろうなと反省しきりで、またしても途中で放り出してしまったが...

理科系出身者は科学全般に強いという大いなる幻想? のおかげもあり、翻訳者としてラッキーなスタートがきれた。小説の登場人物が何を考えているのか、この文章の作者の意図は、などのいわゆる「現代国語」に高校時代あまり関心がもてなかったという単純な理由で理科系に進んだこともあるが、科学的文章の翻訳の方が論理という強い見方があるため取り組みやすい。Native speaker 以外の人が書いた英文は文法的ミスもあり、文章が完結していないものさえあるが、科学的文章の場合は、資料を調べ、文脈をたどればある程度自信をもって著者の意図を推測し、翻訳の段階で文章の完成を手伝えると

いう楽しみも稀にある(間違った楽しみかもしれないが)。著者の独りよがりではなく、論理をたどれば必ず理解できるという手ごたえがある。また、同じ意味を表現するのに、どの言葉、どんな文章を選ぶかにそれほどなやまなくても、事実と論理という強力な助っ人の力をかりて忠実な翻訳文が作れるという心強さもある。

もちろん、科学的文章がかならずしも無味乾燥な あじわいのないものではなく、記憶が違わなければ ポーリング博士(ノーベル賞受賞の化学者)が大学 1年生向けに書いた教科書の英文はとても美しく、 人間性がにじみでるような香り高い文章だったと思 う。同じテーマでも、書き手が違うとまったく趣が 違う。翻訳者も正確さにくわえて香り高い文章が書 ければ一人前になれるのだろうが、むずかしい作業 であることは毎日痛感させられている。

それとは矛盾するが、遠くない将来に機械が大きな力を発揮するようになると思う。特に科学的文章では。ただし、入力する文章はできるだけ短く区切る、意味に曖昧さのない文章であることが必要になろう。これは言葉としての美しさを損なうのでは、という不安もあるが、ボーダレス時代の膨大な情報を言語の壁をこえて活用するには機械翻訳は不可欠であり、そのためには言語そのものを翻訳しやすい構造に変えざるをえない状況が生まれるかもしれない。多少の(この表現には異議があるかも知れないが)犠牲を払っても、充分その意義がある時代になるのかもしれない。

あまりうれしくない予測ではあるが、翻訳者は将 来、機械が理解できるような文章作りを助けるイン ターフェースの役割に甘んじることになるのだろう か。

- 著者の経歴 -

お茶の水女子大学理学部化学科卒業。三菱レーヨンに勤務した後、現在は、フリーランスの翻訳者として、科学、医学・薬学、経済等の翻訳を手がける。

IAMT理事会報告 国際機械翻訳連盟(IAMT)の理事会がサンディエゴで開催されたMT評価ワークショップ時に併せて開催された。会議では長尾会長より連盟事業の総括報告、監査より収支報告が、また各地域機械翻訳協会の活動報告が行われた。

AMTA(アメリカ機械翻訳協会)「機械翻訳に関する研究、評価及び普及を目的とした協会の目的変更、非営利団体としての認可があった」との報告。現在の会員は239件、内12法人2団体が含まれている。会員動向とMT関連活動をまとめた年鑑「MTイエローページ」の発行準備が進められている。

EAM T (欧州機械翻訳協会) 現在の会員数は80件内 7 法人14団体が含まれる。93年は「辞書と機械翻訳」 (4 月ハイデルベルグ)「ブルガリア機械言語学サマースクール」「第2 回評価者フォーラム」の主催を予定。又 MT Summit V の候補地にノリッチ、ブリュッセルが上げられているとのこと。

IAMT編集長(J,Huchins) の指揮のもとに、MT年鑑 (MT Dictionary) の編集が進められており、7月のMT Summit VIに間に合うように刊行される予定。

TMI'93に協賛………自然言語処理学会主催の「TMI'93」が7月14~16日に京都国際交流会館で開催される。(詳細は表2広告欄参照)メインテーマは「次世代の機械翻訳」であり当協会とも密接な関係にあるため協賛、協力する事になった。

利用技術研究会……AAMT利用技術研究会は1月20日に準備会議を開催し、本格的な研究開始のための方向付けを行った。当面システムのプロセス管理、ユーザ辞書の品質確保、文書処理に係わる諸問題の3つの切り口から実態調査(アンケート)を行い、問題点の把握と解決法を研究していく。次回は2月19日開催するが、テクニカルライティングのヒアリングもあわせ実施の予定。

工業英語検定試験……文部省認定の工業英語検定試験が5月23日に実施される。これはテクニカルライターの基礎から実務レベルまでの実力を判定するもので申込締切期限は4月26日、申込先は日本工業英語協会へ(☎03-3434-2350)

VLKB会員募集……(財)京都高度技術研究所では91年度より活動を行っている大規模知識ベース構築プロジェクトの一環として93年度は次ぎの2つのテーマに関してプロジェクト参加企業を募集中です。1知識コミュニティプロジェクト2エキスパートシステム構築支援プロジェクト本件にご関心のある方、参加希望の方は下記へ京都市下京区中堂寺南町17京都リサーチパーク内(財)京都高度技術研究所 VLKB事務局(担当北尾)(☎075-315-8652 FAX 075-315-2898)

EDR成果発表会…… (株)日本電子化辞書研究所では基盤技術研究促進センターの後援を得て3月18日に「ホテルイースト21」(東京都江東区東陽町6-2)で研究成果発表会を開催する。この発表会では「これからの言語処理」と題してAAMT長尾会長が講演を行う他単語辞書、概念辞書、対訳辞書、共起辞書、運用支援システムなどの成果説明やデモンストレーション、パネル討論が行われる。参加費は無料。問合せ先は東京都港区三田1-4-8三田国際アネックス(株)日本電子化辞書研究所成果発表会担当まで

ReadyPenデモ……富士通では英文ビジネスレター作成支援システムReadyPen VO2に加えて光ディスクで動作するコンパクトなMO版を発表したが、(新製品紹介欄参照)テレフォンカード付きアンケートを実施中。またデモンストレーションの希望があれば連絡をしてほしいとの事である。東京都大田区新蒲田1-17-25(情報処理システムラボ)富士通(株)SE情報化推進部(☎03-3737-3235)

(\$\infty\$ 03-3798-5521 FAX 03-3798-5335)

DEFO参加募集……(財)京都高度技術研究所では産学交流の場として「データエンジニアリングフォーラム」(略称DEFO)を92年6月より開催、オブジェクト指向データベース等新しい研究分野のチュートリアルやデータモデルの標準化問題等の意見交換を行っている。現在93年度の法人会員(年会費20万)を募集中。問合せ先は上記の所へ。

委員会活動状況報告

●運営委員会 11月13日 臨時理事会開催結果報告 海外会員募集について 各委員会報告等

12月10日 IAMT会則の確認 日経新聞出稿について 翻訳フエア準備状況報告等

2月3日 TMI協賛名義 TMI協賛金支出 IAMT収支報告び会費支払いについて他

●翻訳フエア実行委員会(企画広報委員会)

11月 5日 展示会小間割抽選 展示要綱説明

11月 9日 開催案内等印刷物内容検討 講師依頼状況確認

11月27日 会場準備状況確認 看板等展示物確認検討

12月 1日 参加申込状況確認 招待基準検討 司会者等決定

●システム評価研究会

11月26日 翻訳困難例文の云い換え例研究 翻訳困難文の分類

2月 4日 翻訳困難例文の分類と取りまとめ 今後の方向づけ

●システム利用技術研究会

12月 3日 研究会の方向付け検討 WG幹事候補の検討

1月20日 研究方向の検討 幹事の役割分担

●ジャーナル編集委員会

11月27日 英文版発行対策 AAMT2号編集方針 IAMTニューズレター対応他

2月 5日 投稿規定、査読規定検討 AAMT3号編集方針 機関誌方針の確認

●MT-Summit展示委員会

12月10日 展示要綱の検討 出展要請社のリストアップ 海外出展社の要請、検討

●MT-Summit広報委員会

1月19日 参加動員策の検討 動員依頼状の検討 協賛依頼先の確認

●MT-Summitプログラム委員会

11月10日 プログラムの検討 講師依頼 司会者の検討

2月 1日 プログラムの修正 講師引き受けの確認 プロシーディングの検討

●MT-Summit実行委員会

11月13日 収支予算案の検討 補助金の見通し 参加予定者の見積

12月10日 収支予算の検討 参加料の検討

会員の住所変更

(株)十 印 (新住所)〒105 東京都港区芝1-12-8 十印ビル(MT部1 03-3455-8716 FAX03-3455-8364)

新会員の紹介

KIM GIL CHANG

JONG-HYEOK LEE

KHE-YIHSU

WULINGFEI

片田江久男

森 誠

小川浩平

佐々木由紀子

本鄉保夫

伊東昌高

第4回機械翻訳サミット プログラム (TENTATIVE)

7月20日(火)

10:00 開 会 式 開会挨拶 長尾 真(IAMT会長、サミット組織委員長)

挨 拶

通商産業省 機械情報産業局

10:15 基調講演

長尾 真(京都大学教授)

11:00 招待講演「機械翻訳技術の最新の動向」

J.Hutchins (英・イーストアングリア大)

12:00 昼 食

13:30 ビッグプロジェクト「多言語機械翻訳」 (財)国際情報化協力センタ-

15:30 パネルディスカッション「機械翻訳技術の将来」

座長 内田裕士 (株)富士通研究所

18:30 歓迎レセプション

7月21日(水)

9:00 ビッグプロジェクト「自動翻訳電話」

轉松 明(株)ATR自動翻訳電話研究所

9:40 ビッグプロジェクト [Verb Mobil]

W.Wahister (独・ザルブリッケン大)

10:20 コーヒープレーク

11:00 ビッグプロジェクト「EC Language Project」

L.Rolling (EC委員会代表部)

12:00 昼

13:30 招待講演「機械翻訳利用技術の現状」

M. Vasconcellos (米・機械翻訳コンサルタント)

14:30 コーヒプレーク

15:00 パネルディスカッション「機械翻訳の評価法」 座長 M.King (スイス・ジュネープ大)

18:30 バンケット(有料)

7月22日(木)

9:00 ビッグプロジェクト コーパスプロジェクト

M.Liberman (米・ペンシルバニア大)

9:40 ビッグプロジェクト「電子化辞書」

(株)日本電子化辞書研究所

12:00 昼 食

13:30 招待講演「電子化辞書から大規模知識ベースへ」

横井俊夫(株)日本電子化辞書研究所

14:30 コーヒーブレーク

14:45 パネルディスカッション「国際協力」

座長 田中穂積 (東工大)

16:30 閉 会 式 次回開催地の決定

サミット宣言

AAMTジャーナルNo.2

発行月 1993年2月

編集委員長 野村 浩郷 (九工大)

発 行 アジア太平洋機械翻訳協会

編集委員 亀井真一郎 杉山健司

所 在 地 〒106 東京都港区赤坂7-2-17-305

永野 文美 宮城 雅之

信田 惠壱 鳴海 武司

電 話 03-3479-4396/4398

砂川昌順

F A X 03-3479-4895

機械翻訳サミット Ⅳ

Machine Translation Summit IV

開催期間 平成5年7月19日(月)~22日(木) 開催場所 ホテルオークラ神戸

(神戸市中央区波止場町1-2メリケンパーク内)

開催日程 7月20日~22日 根

7月20日~22日 機械翻訳サミット(詳細は24頁参照)

7月19日10~14時 テクニカルツアー「ATR自動翻訳電話研究所」見学

7月19日14~18時 ワークショップ 「シソーラス」

7月19日14~18時 チュートリアル「機械翻訳の使用法」

7月19日18:30~ ウエルカムレセプション

7月20日18:30~ バンケット(有料)

7月19日~22日 機械翻訳システム及び関連商品展示会

参加料金 機械翻訳サミット

4月末日迄に申込み入金の場合

30,000円

5月1日以降の申込み入金

40,000円

AAMT.JEIDA会員は、開催日まで30,000円と期間メリット適用

ワークショップ、チュートリアル

各10,000円

バンケット

15,000円

テクニカルツアー・ウエルカムレセプション,展示会は入場無料

参加人員 機械翻訳サミット

300名

ワークショップ,チュートリアル

各100名

テクニカルツアー

(バス乗車定員有り)

(満員になり次第締切る事があります)

主 催 アジア太平洋機械翻訳協会(A A M T) (社)日本電子工業振興協会(J E I D A)

後 援(予定) 通商産業省(MITI)

協 賛 (予定) 情報処理振興事業協会 (財)情報処理事業開発協会

(財)国際情報化協力センター (財)大阪科学技術センター (財)関西情報センター (社)情報処理学会 人口知能学会 計量国語学会 (社)電子通信学会 自然言語処理学会 日本ソフトウェア科学会 日本認知科学会 他関係団体

