

MT News

International

Newsletter of the International Association for Machine Translation

ISSN 0965-5476

Issue no.1 January 1992

In this issue:

- * **Association news (IAMT, AMTA, EAMT, JAMT)**
 - * **Conference reports**
 - * **Programmes, policies, surveys**
 - * **New products**
 - * **Research developments**
 - * **Recent publications**
 - * **Conference announcement (TMI-92)**
 - * **Forthcoming events**
-

Editor-in-Chief:

John Hutchins, The Library, University of East Anglia, Norwich NR4 7TJ, United Kingdom.

Fax: +44 (603) 259490; Email: L101@uk.ac.uea

Assistant editors:

AMTA: Dr. Joseph E. Pentheroudakis, Executive Communication Systems, Inc., 455 N. University, Suite 202, Provo, Utah 84601, USA. Fax: +1 (801) 374-6292; Email: pentherj@edu.utah.cc, or: ecs@yvox.byu.edu

EAMT: Dr. Tom C. Gerhardt, CRP-CU/CRETA, 13 rue de Bragance, L-1255 Luxembourg.

Fax: +352 (44) 73 52; Email: tom@crpculu.uucp

JAMT: Professor Hirosato Nomura, Kyushu Institute of Technology, Iizuka, 820 Japan.

Fax: +81 (948) 29-760 ; Email: nomura@dumbo.ai.kyutech.ac.jp

Publisher: Praetorius Ltd., 5 East Circus Street, Nottingham NG1 5AH, UK.

From the Editor

This newsletter has been founded for the exchange of information and opinions about machine translation (MT). It is open to all with any interest in this increasingly important field: users (actual and potential), manufacturers and vendors, sponsors and supporters, researchers and developers, and any others who want to know what is going on and what the future may bring. This first issue is necessarily larger than later ones are expected to be, primarily because it carries essential information about the foundation of the International Association for Machine Translation (IAMT) and the regional associations, AMTA, EAMT and JAMT. But it is larger also because it reports on MT activity for much of the past year.

It is intended that MT News International will appear initially three times a year in the months of January, May and September. It will carry regular features reflecting the varied interests and approaches of the expected readership: news of activities by the associations, reports of new products from manufacturers, research developments, new installations of systems, users' comments and opinions, and reports of conferences, seminars and workshops devoted to or related to MT. It will inform readers about programmes, innovations, policies, financial support and sponsorships by and for national and international (governmental and non-governmental)

organizations and institutions. It will report on surveys of MT activity, evaluations of systems, and assessments of the MT market, and of course it will include notices of recent publications (books, journals and reports) and a calendar of forthcoming events. It is hoped that a lively correspondence section will develop, where in particular users of systems (regular, occasional or just potential, and whether translators or non-translators, companies or individuals) will express their views and needs, hopes and disappointments.

A major goal of IAMT and of *MT News International* is to provide access to impartial information about systems. It will seek to maintain a balance of views and opinions and not become the exclusive organ for any one particular interest or persuasion. Above all, it is hoped that this newsletter will enable open and frank communication between all those with an interest in MT in all its various forms.

Contributions to *MT News International* should be sent to the Assistant Editors in the regional associations or directly to the Editor-in-Chief at the addresses given on the left (and whenever possible in electronic form). Items for inclusion in the section on 'Recent publications' should be sent to the Editor-in-Chief, who thanks all those who have already responded to requests. He is also most grateful to Sarah Mortimer of Mitaka (Leamington Spa, UK) for translating excerpts from the JAMT Journal.

Copies of the newsletter are distributed by each regional association to their members, and enquiries about obtaining copies should be addressed to the relevant secretariats (as given elsewhere in this issue). At some future date it is hoped that the newsletter will be available in electronic form on an appropriate network as well as appearing in its present printed form. Comments and suggestions on any aspect or feature of *MT News International* will be welcomed by the editors.

From the President of IAMT

Makoto Nagao

The International Association for Machine Translation (IAMT) was established last July on the occasion of the Third Machine Translation Summit Meeting (MT Summit III) in Washington D.C. It aims at bringing together users, developers, researchers, sponsors, and other individuals or institutional or corporate entities interested in machine translation (MT) for the purpose of promoting and fostering, by every available means, the development and active use of MT systems. The IAMT offers opportunities and occasions to exchange information, study MT technologies and applications, and discuss and establish reference criteria or standards in areas of common interest to its members.

In 1987 the first MT Summit meeting was held at Hakone in Japan, where it was clearly recognized that the exchange of experience, information and opinions among users, developers, researchers, and government people who support the activities of MT is of great value for the development and wider use of MT systems. From this experience we organized the International Forum on MT at Oiso in Japan in 1989, the MT Summit II in Munich in 1989, and the MT Summit III in Washington D.C. in 1991. Those who attended these conferences expressed the wish to have an association for the exchange of information and for discussing the various problems of MT. The idea was discussed at MT Summit II, and finally at MT Summit III we established IAMT and at the same time three regional associations: the European Association for Machine Translation (EAMT), the Association for Machine Translation in the Americas (AMTA), and the Japan Association for Machine Translation (JAMT). JAMT is making an effort to extend its activities to other Asian countries and to form an Asian Association for Machine Translation in the near future.

The activities of IAMT will include the publication of newsletters, the organization of the MT Summits, workshops on a variety of topics, symposia, tutorials and others events, proposals

for standards, and it will serve as a clearinghouse for current information on MT. This MT News International is the first issue of the IAMT newsletter. Since the IAMT was established only a few months ago, we have still no significant activities of our own, but we are endeavouring to give good service to members of the association. It has already been decided that the Fourth MT Summit will take place in Kobe, Japan, in July 1993; and we are planning the formation of some technical committees, such as one on the evaluation of MT systems. We shall welcome any ideas, proposals and comments you may have for the promotion of IAMT activity. I hope that our activities will be valuable to all members, and that all people with any interest in the field of MT will join the association, whether users, developers, manufacturers, vendors, researchers, translators and any others from whatever background who wish to contribute to the furtherance of machine translation.

MT Summit Sees IAMT Kickoff

Muriel Vasconcellos

The first general membership meeting of the International Association for Machine Translation, held on 4 July 1991 as part of the final day's programme of MT Summit III, represented the culmination of more than two years of preparatory activity.

The idea of an international association bringing together people interested in all aspects of MT was first voiced in April 1989 at the International Forum on Translation Technology in Oiso, Japan, and was reiterated by Professor Makoto Nagao in his keynote address later that same year at MT Summit II in Munich. Shortly thereafter a prospectus setting forth the Association's purpose and the basic principles by which it would be guided was drafted by Makoto Nagao working in collaboration with Veronica Lawson and Muriel Vasconcellos.

The challenge was to form an institution that would be truly international and at the same time strong and active at the local level. The Drafting Committee felt that this dual aim could be accomplished most effectively through the simultaneous creation of three regional associations – in the Americas, Asia, and Europe – which would be woven, in turn, into a worldwide structure that would be tied to the biennial MT Summits. At all levels the work of the Associations would give equal honour to MT researchers, commercial developers, and current or prospective users. Above all, the Associations would be democratic. The prospectus incorporating these principles was announced and circulated in Washington, D.C., on 7 December 1989 at the Symposium on Japanese-to-English Machine Translation.

Over the next 18 months initiatives got under way in the Americas, Europe, and Japan--the last of which is now being coordinated with a pan-Asian group. Thus, by the time MT Summit III rolled around in July 1991, each of the three regional associations was in place and each had been able to hold its first meetings and elect its respective members of the international Council.

At the first meeting of the Council, held on 3 July, Makoto Nagao was elected president; Margaret King, vice-president and president-elect; Muriel Vasconcellos, secretary; and Roberta Merchant, treasurer. John Hutchins was appointed editor-in-chief of the newsletter; he will collaborate with regional editors Joseph Pentheroudakis (Americas) and Tom Gerhardt (Europe) and be assisted by Geoffrey Kingscott. It was agreed that the work of the IAMT secretariat will be divided between the regions: the Americas will provide administrative support, with IAMT to be registered as a corporation in Washington, D.C.; Europe will be responsible for information, including both the newsletter and a clearinghouse; and Asia will coordinate planning and programming. Membership in IAMT will normally be through the regional associations, each of which will contribute 10% of its dues to the international treasury. English is the official language of IAMT. It was decided to make the newsletter available in electronic form to the three regions, which in turn will assume responsibility for hard-copy dissemination.

Against this background, the general membership meeting was called to order by Professor Nagao at 2:15 p.m. on 4 July. Secretary Muriel Vasconcellos reviewed the main decisions of the Council and summarized the highlights of the American and European initiatives to date. Whereupon Makoto Nagao declared that IAMT had officially come into being.

ASSOCIATION NEWS

Summary of the IAMT constitution and officers.

The charter of IAMT envisages the following activities:

- (a) the collection and compilation of information by serving as a repository for historical and current documentation on MT, converting the texts to machine-readable form when and as feasible.
- (b) the exchange of information by serving as a clearinghouse for current information of interest to MT users and developers, e.g. for users: translation market trends, available MT systems, types of MT applications, introduction of MT systems, approaches to text file input, pre- and postediting strategies, evaluation of MT, etc.; and for developers: theories of MT, MT technologies, improvement of MT, approaches to dictionary-building, data bases and other files available for exchange, etc.
- (c) the dissemination of information, by publication of a regular Newsletter on MT, an MT handbook (including agreed definitions of terminology) and bibliographies
- (d) the development of reference criteria and standards in such areas as: common document format for MT input, exchange format for dictionaries, design of controlled language, evaluation of MT output and MT systems

The IAMT has three categories of membership: individual, corporate, and institutional. There are at present three Regional Associations (for the Americas, Europe and Japan), whose members are automatically members of IAMT. Regional Associations have their own boards and carry out their own activities in addition to participating in those of the IAMT. The supreme governing body of IAMT is its General Assembly, convened in conjunction with the MT Summit. The IAMT Council executes decisions of the General Assembly and carries out IAMT business; it will have ten members, three representatives from each regional association and the Editor-in-Chief of the newsletter. From its ranks it will elect a President, a Vice President, a Secretary, and a Treasurer. The office of President will rotate every two years between the three Regional Associations. The Editorial Board of the IAMT Newsletter is composed of three editors, one from each Regional Association designated by the respective Regional Board, plus the Editor-in-Chief designated by the IAMT Council.

The IAMT will convene the MT Summit every two years, rotating between the three regions. The Association will also: sponsor and support workshops, symposia, and conferences on MT and related technologies and applications, organize tutorials and training courses on MT applications and skills involved in the use of MT and on MT technologies, and establish technical committees, special interest groups, and study teams.

Officers of IAMT:

President: Professor Makoto Nagao (IAMT)
Vice President: Dr. Margaret King (EAMT)
Secretary: Dr. Muriel Vasconcellos (AMTA)
Treasurer: Roberta Merchant
Editor-in-Chief: John Hutchins

ASSOCIATION FOR MACHINE TRANSLATION IN THE AMERICAS

[The following is an extended excerpt from an open letter by Muriel Vasconcellos, AMTA president, included in the Association's first membership drive]

Dear Colleague:

I am writing you to invite you to join the new Association for Machine Translation in the Americas (AMTA).

This initiative is now a reality, and it is an idea whose time has come. Machine translation is at long last beginning to come into its own. The signs are subtle, but they are there: at MT Summit III last July, the developers and vendors of 22 working systems crowded into the Exhibit Hall and made their products do what they were supposed to do in the presence of demanding onlookers. Moreover, 11 systems were demonstrated live in a theatre-style setting. A number of them were submitted to the acid test of translating random input under public scrutiny. The fact that so many developers chose to take part in this event is testimony to the incrementally changing times: dictionaries are larger and more deeply coded, parsers are making fewer mistakes, hardware and programming languages are turning smarter and more powerful, fresh approaches are coming out of the laboratory and moving into the mainstream, and, as a result of all this, MT output is gradually becoming more reliable.

At the same time, interest in machine translation has never been greater, as the planet's economy and its safe-keeping become increasingly dependent on information. In a world where national boundaries are getting erased, information must be translated into other languages in order to sell products, advance research, maximize the earth's bounty, keep people wise and healthy, and stop wars.

Now more than ever there is need for an association devoted exclusively to machine translation – one free of other competing concerns that could concentrate its energies totally on fostering this technology. AMTA was incorporated last April in Washington, D.C., and held its first organizational meeting at the time of MT Summit III. Its purpose is to:

...bring together researchers, commercial developers, users, sponsors, and other individuals or institutional or corporate entities interested in machine translation in order to promote and foster the development and active use of machine translation systems.

In joining together individuals and institutions that approach the MT challenge from different perspectives--representatives of academia, the marketplace, and government; researchers and commercial developers; users, would-be users, and just plain watchers – AMTA's intention is to encourage new relationships, create a climate for the exchange of ideas, and give importance to all points of view. This commitment is reflected in the composition of the initial Board of Directors and in the Board's determination to perpetuate a balanced representation.

AMTA's effort in the Americas is being replicated in Europe and Asia, and the three regional associations are banding together to form the International Association for Machine Translation (IAMT), which also held its first meeting at the time of MT Summit III.

You are invited to join us in this pioneering initiative. AMTA, together with IAMT, is already looking forward to a full plate of activities. The first issue of a regular newsletter on MT activities worldwide under the able editorship of John Hutchins is scheduled to appear in January 1992. In addition, AMTA will be organizing workshops and conferences, co-sponsoring the biennial MT Summit, proposing criteria for the evaluation of MT systems, and cooperating in the exchange of "pre-competitive" data such as parallel texts and mailing lists – to mention but a few of the projects in the pipeline. Happenings planned for 1992 include a "buyers' workshop" to

showcase MT systems to potential purchasers and a tutorial on MT to be offered in cooperation with the American Translators Association.

Exciting times are ahead. I hope that you will be with us as we embark on this new undertaking.

Sincerely,

Muriel Vasconcellos
President

Activities of AMTA [from the Constitution]:

- "a. Encourage and advance MT technologies and suitable applications thereof;
- b. Share knowledge that contributes to the development and active use of MT systems, which may include the collection, compilation, exchange, and dissemination of information;
- c. Sponsor and support workshops, symposia, and conferences on MT and related technologies and applications;
- d. Develop appropriate training materials and programs;
- e. Reward excellence in machine translation;
- f. Facilitate access by researchers to machine-readable corpora and cooperate in the promotion of exchange formats and text-encoding conventions;
- g. Help to set priorities for MT development, avoid unnecessary duplication of effort, and provide guidance for investments in areas of highest priority; and
- h. Discuss and establish reference criteria or standards in areas of common interest to its members.

AMTA registered office: 655 Fifteenth Street, N.W., Suite 310, Washington, D.C. 20005

AMTA Officers and initial Board of Directors:

President: Muriel Vasconcellos (Pan American Health Organization, Washington, DC).
Tel: (202) 861-4338; Fax: (202) 223-5971, (202) 667-8808

Vice-president: Sergei Nirenburg (Carnegie Mellon University, Pittsburgh, PA). Tel:
(412) 268-6593; Fax: (412) 268-6298

Secretary: Winfield Scott Bennett (Linguistics Research Center, University of Texas,
Austin, TX) Tel: (512) 471-4566; Fax: (512) 471-9646

Treasurer: Roberta Merchant (U.S. Department of Defense). Tel: (301) 688-6149

Newsletter editor: Joseph Pentheroudakis (Executive Communication Systems, Inc., Provo,
UT). Tel: (801) 377-1167; Fax: (801) 374-6292

Directors: Rob Billingsley (Defense Tech. Information Center),

Dale Bostad (Foreign Technology Division, US Air Force),

Robert M. Carswell, Jr. (Counsel),

Tony Centodocati (National Science Foundation),

John Chandiox (John Chandiox Experts-Conseils),

Mark Eaton (Microelectronics and Computer Technology Corporation),

Denis Gachot (Systran),

Deanna Hammond (American Translators Association),

Bernard Scott (Logos Corporation),

Tom Seal (ALPNET),

Howard Teicher (Translation Technologies International),

Virginia Teller (Hunter College),

Ming Tsui (Canadian Secretary of State),

Yorick Wilks (New Mexico State University)
Committees and Committee Chairs - Bylaws: Scott Bennett/Mark Eaton;
Education: Deanna Hammon/Doris Cabezas;
Ethics: Howard Teicher/Ken Goodman;
Evaluation: Sergei Nirenburg/Bernard Scott;
Fundraising: Howard Teicher;
Information: Rob Billingsley/John Chandioux;
Membership: Joann Ryan/Scott Bennett;
Nominations: Joseph Pentheroudakis/Virginia Teller;
Program: Rob Billingsley/Denis Gachot

AMTA Holds First Meeting of its Initial Board

Scott Bennett and Muriel Vasconcellos

Fifteen of the 18 members of the Initial Board of Directors of the Association for Machine Translation in the Americas (AMTA) met in Washington, D.C., on July 1, 1991, immediately prior to MT Summit III to conduct the Association's first business. Those present were: Scott Bennett (Siemens-Nixdorf); Rob Billingsley (U.S. Department of Defense); Robert Carswell (counsel); Tony Centodocati (U.S. National Science Foundation); John Chandioux (John Chandioux Experts-Conseils); Mark Eaton (MCC); Denis Gachot (SYSTRAN); Roberta Merchant (U.S. Department of Defense); Sergei Nirenburg (Carnegie Mellon University); Joseph Pentheroudakis (ECS); Bernard Scott (LOGOS); Howard Teicher (TOVNA); Virginia Teller (Hunter College); Ming Tsui (Canadian Secretary of State); and Muriel Vasconcellos (Pan American Health Organization).

After approving a set of provisional bylaws, the Board elected its officers, as follows: Muriel Vasconcellos, president; Sergei Nirenburg, first vice president; Scott Bennett, secretary; and Roberta Merchant, treasurer. In addition, Robert Carswell was appointed vice president for legal affairs and general counsel, and Joseph Pentheroudakis was appointed editor, to coordinate AMTA contributions to the international newsletter.

The dues structure was decided on as follows: active members, US\$65.00 (\$50.00 for those joining prior to 1 January 1992); associate members, US\$25.00; institutional (not-for-profit) members, US\$200.00; and corporate (for-profit) members, \$400.00.

Mandates, chairpersons, and initial membership were agreed on for the following committees: Bylaws, Ethics, Evaluation, Information, Nominations, and Program (as listed above).

At a general meeting of AMTA members, held in the late afternoon of 4 July, further details of the Association's work were hammered out and the corresponding committees were expanded.

EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION

[from a circular sent to potential members]

The European Association for Machine Translation (EAMT) is a newly formed organization which seeks to bring together researchers, developers, users, sponsors, and other individuals or institutions interested in machine translation and/or computer assisted translation whatever their background. It is one of three regional associations of the International Association for Machine Translation, which was set up officially in July 1991. The aim of EAMT is to promote and foster the development and active use of machine translation and/or computer assisted translation, and to provide a forum for the exchange of information and ideas among all those interested and involved in this technology. The association is intended to represent the views of the machine translation community as a whole, and to be the major sources for objective information about

machine translation. Together with the other regional associations, EAMT will encourage research of all kinds in the field, promote the exchange of information, seek to establish reliable methods of evaluation, organize seminars, lectures, courses and workshops for users and potential purchasers of systems, and disseminate information on conferences and publications concerned with machine translation, its users and its applications.

The EAMT membership rates for 1992 will be: Individual members SFr. 35.-; Non-profit making institutions SFr. 175.-; Profit making institutions SFr. 350.-

All EAMT members are automatically members of the International Association for Machine Translation, and will receive the IAMT newsletter *MT News International*.

EAMT aims and activities

[from the Articles of Association]

"The Association's purpose shall be to bring together users, developers, researchers, sponsors and others interested in the field of machine and/or computer assisted translation research, development and use.

In the pursuit of the above mentioned purpose, the Association shall in particular perform the following activities:

- promote the exchange of information on machine and/or computer assisted translation among its members.
- encourage the development of computer software in the field of machine and/or computer assisted translation.
- encourage development of models and theories related to machine and/or computer assisted translation.
- organize seminars and lectures in the field of machine and/or computer assisted translation.
- disseminate information on machine translation workshops and conferences, as well as on publications relevant to the domain.
- participate in and/or become a member of other Associations that pursue the same or similar purpose."

The supreme organ of EAMT is the General Assembly of all its members, which elects the President and members of the Executive Committee. Voting rights reflect different categories of membership (individual members have one vote, non-profit making institutions five, and profit making institutions ten). The Executive Committee is composed of the EAMT President and four members elected by the General Assembly; additional members may be co-opted with approval of the General Assembly.

A provisional executive committee was formed on 4 July 1991 with the following officers and responsibilities:

- President: Margaret King (ISSCO).
- Secretary: Ian Johnson (Sharp Laboratories of Europe Ltd.)
- Treasurer: Doris Albisser (Union Bank of Switzerland, Zu"rich)
- Newsletter assistant editor: Tom C.Gerhardt (CRP-CU/CRETA)
- International activities: Ulrich Heid (University of Stuttgart)
- Lexicography: John Hutchins (University of East Anglia)
- Education: Kai Lee Hong
- PR activities: Karin Meichtry (Ascom Hasler Ltd.)
- New activities/fund-raising: Brigitte Roudaud (B'VITAL/SITE)
- IAMT representative: Ulrike Schwall (IBM Deutschland)
- Newsletter representative: Sylvie Wallez (Universite' Lie\ge)
- Secretariat: Tamara Wehrli

Address of the EAMT Secretariat: ISSCO, 54 route des Acacias, CH-1227 Carouge, Geneva, Switzerland

Workshops planned by EAMT:

Machine Translation and Translation Theory, July 1992, Berlin (Germany). For details see 'Forthcoming events' at the end of the Newsletter.

MT & Evaluation, early October 1992, Geneva (Switzerland).

MT & Lexicon, October 1992, Heidelberg (Germany).

MT & MAT and its Users, end of 1992, Spain.

For further information please contact: Ulrike Schwall, IBM Deutschland, Institute for Knowledge Based Systems, Computational Linguistics 2, Tiergartenstrasse 15, Postfach 10 30 68, D-6900 Heidelberg, Germany

JAPAN ASSOCIATION FOR MACHINE TRANSLATION

The Japan Association for Machine Translation (JAMT) was established on April 17th, 1991, and held its first general meeting on 25 July 1991.

Aims and activities

[from the constitution]

"The purpose of JAMT is to contribute its every effort to the development of the machine translation systems in the world by (i) establishing a common ground for the mutual exchange of any relevant information and the study and discussion of any subject common to the systems among manufacturers, users and research institutions interested in such systems, (ii) cooperating with any academic society or institution relating to such systems in the world, and (iii) extending its business activities to foreign countries, so as to realize the steady advancement and dispersal of such systems in society.

- JAMT shall attempt to organize an Asian association for machine translation to promote the development of the machine translation systems in Asian countries

- JAMT shall engage in the following...

(1) Holding of various meetings for research, lectures, discussions, training, educational visits, etc.

(2) Publication of newsletters and any other forms of literature, whether or not periodical.

(3) Education of users for MT systems, establishment of a training course for such users, and the conducting of tests for measuring technical skill in such systems.

(4) Technical development of MT systems.

(5) Establishment of the standard and specifications regarding MT systems, especially the preparation of any recommendation concerning the evaluation method...

(6) Preparation of any guideline for the introduction and use of MT systems.

(7) Technical investigation and the collection and exchange of any information or literature.

(8) Promotion of any academic or scientific investigation and study.

(9) Establishment of any cooperative activity with any association society, institution or organization of similar kind inside or outside Japan.

(10) Establishment of any business other than mentioned in the preceding paragraphs necessary for achieving the purpose of JAMT."

The supreme body of JAMT is the General Assembly of its members. Present officers and board members are:

President: Makoto Nagao (Kyoto University)

Vice-presidents: Kunika Mizushima (Toshiba Corporation), Taizo Kotani (Intergroup Co.Ltd.)

Directors: Atsushi Asada (Sharp Corp.), Eiichi Ohno (Mitsubishi Electric Corp.), Tatsuya Komatsu (Simul International Co.Ltd.), Takeshi Suzuki (JEIDA), Hozumi Tanaka (Tokyo Inst.Technology), Masao Toka (NEC Corp.), Yasukazu Toda (CSK Co.Ltd.), Masaya Nakajima (Matsushita Electric Industrial Co.Ltd.), Takashi Nonouchi (Hitachi Ltd.), Hirosato Nomura (Kyushu Inst. Technology), Katsuhiko Hiura (Jujirushi Co.Ltd.), Toshio Hiraguri (Fujitsu Ltd.), Masataka Yamamoto (Oki Electric Industry Co.Ltd.)

Auditors: Mihoko Katsuta (Japanese Translation Federation Inc.), Seishun Sato (JEIDA)
Committee chairmen

- Steering Committee: Hiroshi Uchida (Fujitsu)
- Newsletter: Hirosato Nomura
- PR Planning: Hiroshi Uchida
- System evaluation: Hozumi Tanaka
- Finance: Akira Furusawa

Address of the secretariat:

Mrs. Megumi Okita, Japan Association for Machine Translation, 305 Akasaka Chuo Mansion, 2-17, Akasaka 7-chome, Minato-ku, Tokyo 107 Tel: +81-3-3479-4396. Fax: +81-3-3479-4895

JAMT fees:

Personal member -- 1,000 yen for entry; 5,000 yen/year

Group member -- 10,000 yen or more for entry; 50,000 yen/year

Organization (MT manufacturer) -- 100,000 yen or more for entry; 500,000 yen/year.

Foundation members:

MT manufacturers: 100,000 yen for set up; 800,000 yen/year.

MT system users: 50,000 for set up; 400,000 yen/year

JAMT Journal

The committee for the JAMT journal began in April 1991 with the foundation of JAMT itself, with Hirosato Nomura as chairman. There are eight members, three from large translation companies, four from large computer companies and one (the chairman) from a university. It is intended that the JAMT Journal will include: essays on translation and machine translation, introductory articles on MT/NLP technologies, essays on MT experiences from individual users, from organizations, and from translators, introductions to MT/NLP research/development institutions, panel discussions on Translation/MT from the viewpoints of research, development, market, administration, etc., news on relevant activities, information on conferences/meetings, news of MT related projects from over the world, information on new MT products, views from JAMT member, etc. JAMT Journal appears every other month, with the first issue appearing on 25 July 1991. Contents of the first three issues are listed in 'Recent publications'.

CONFERENCE REPORTS

MT Summit III in Washington, July 1991

John Hutchins

The third MT Summit took place from 1st to 4th July 1991 in the elegant surroundings of the Mayflower Hotel, Washington, D.C. The conference was ensured of success by the excellent organization of Sergei Nirenburg and his colleagues from the Center for Machine Translation of

Carnegie-Mellon University, Pittsburgh, with local arrangements in the capable hands of Muriel Vasconcellos. Nearly 400 registered for the conference, with a particularly strong participation from Japan.

The first day was devoted to 'executive briefings' (intended primarily for participants unfamiliar with the field.) Harold Somers and John Hutchins provided a general overview of the basic issues in MT research, including a brief history, an outline of basic system design, practical aspects (pre- and post-editing, interactive use, etc.), and the major linguistic aspects of analysis and representation, ambiguity, contrastive lexical and grammatical problems and stylistic considerations. They were followed by a joint presentation by Zenshiro Kawasaki and Joann Ryan on how MT can be successfully integrated into document production services of organizations. The various options were covered, from submitting disks to a service bureau utilizing MT to purchasing or leasing a system for internal use, with guidelines for selecting appropriate systems and for cost-effective utilization. In the afternoon, Jaime Carbonell and Sergei Nirenburg described the knowledge-based approach to MT system design, dealing with the kind of knowledge required for translation and the necessary computational architecture, processing algorithms, methods of syntactic and semantic analysis and target text generation. They were followed by Makoto Nagao describing the current situation in Japan, the systems developed and available, and the practical use of systems.

The conference proper opened on 2nd July with speeches from Jaime Carbonell (general chairman) and Makoto Nagao, followed by accounts of Eurotra by Jörg Schütz and others and of the LMT project at IBM by Mori Rimón and others and two panel discussions. The first devoted to 'the MT user experience' covered recent developments in Meteo (John Chandiooux), at PAHO (Muriel Vasconcellos), at Logos (Bernard Scott), at the Japan Broadcasting Corporation (NHK) in the use of systems for subtitles and news services (Hideki Tanaka), and at the US Defense Technical Information Center using Systran (Rob Billingsley). The second panel was devoted to 'building the customer base', with presentations by Michel Gainet (United Nations), Muriel Jerome-O'Keeffe (CACI), Edith Losa (Stromberg-Carlson), Maria Martínez-Pérez (IBM Spain), Sue Walker-Toledo (Netrologic), and Michael Zarechnak (Georgetown University). The afternoon of the first day concluded with papers on the Ultra experiment at New Mexico State University, research on semantics for interlingua-based MT at the Microelectronics and Computer Technology Corporation, and the corpus- and statistics-based ArchTran system for English-Chinese translation under development in Taiwan.

The second day continued with presentations of research systems and projects and a further three panel sessions. Many of the presentations were given in parallel sessions, presenting difficult choices for conference participants. They covered developments of existing commercial systems (Siemens' METAL, NEC's PIVOT, Toshiba's AS-TRANSAC), and a wide range and variety of experimental projects: translation of avalanche bulletins at ISSCO, bilingual lexica at Cambridge University, the ALT J/E under development at NTT, the KIELIKONE translation workstation for Finnish, the translator's workbench project at Triumph-Adler, another paper on the Ultra project and three further presentations of the research work at Carnegie-Mellon University (the knowledge-based MT system KANT, 'memory-based' MT for the SNAP parallel processor, and the JANUS speech translation system). A marked feature was the large number of systems involving Japanese and English, under development in both Japan and the United States.

The three panel sessions on the second day were devoted to international perspectives, the respective roles of translators and MT, and the evaluation of systems. In the first panel session, two speakers reported on the changing attitudes to MT in the United States (Deanna Hammond from the perspective of the American Translators' Association and from her experience in US federal agencies, and Clark Wayne from the viewpoint of the US Department of Defense). There are clearly strong prospects of more substantial support for MT research than in the recent past, with an emphasis on the exploration of the newer AI, statistical and corpus-based approaches, on

integration with other text handling systems, and on realistic evaluations of performance. Nick Ostler summarized the UK position, emphasising the importance of the 'language industries' market in the European Communities and the perhaps relatively smaller role of MT research in this context. Jan-Michel Czermak concentrated on the important contribution of German research in the Eurotra project and summarized the findings of the SCS study (reported elsewhere in this issue). The two Japanese participants were Watanabe, who reported on activity at the Japan Information Center of Science and Technology (particularly the development of the Mu system), and Makoto Nagao, who spoke about the multi-national CICC project (involving Japan, China, Thailand and Indonesia).

The second panel of the day was entitled 'Where do translators fit into Machine Translation?' Claude Bédard argued for more realistic tools for translators, which actually save them time and effort; Harold Hille (United Nations) advocated closer collaboration between translators and researchers, while not abandoning the ultimate FAHQT goal; Martin Kay (Xerox Corporation) was doubtful about the benefits of present systems, believing that they were more often hindrances than aids for translators; Fred Klein maintained that MT developers had ignored the real needs of practising translators and that extravagant and false claims were antagonising potential supporters; Sergei Nirenburg (Carnegie-Mellon University) also advocated greater involvement of translators in MT research of all kinds; and Pierre Isabelle wanted MT research to tackle what human translators found difficult and not what they found easy (such as pronouns and anaphora). A common theme of all presenters was that MT researchers should aim to satisfy real needs and not be concerned with 'interesting' academic problems. The panel was chaired by Alex Gross who proposed a code of ethics to which MT developers should subscribe.

The third panel was devoted to the 'evaluation of MT systems' chaired by Margaret King (ISSCO) who introduced the session with a succinct overview of the factors involved, based in large part on discussions at the Evaluators' Forum in April 1991 (reported elsewhere in this issue.) Yorick Wilks (New Mexico State University) was highly skeptical of the usefulness of evaluation, either to researchers and developers or to potential purchasers; he was particularly doubtful about the very possibility of independent 'test suites' as opposed to evaluations in specific environments with real texts. Sture Allen (Gothenburg University) and Ulrich Heid (Stuttgart University) outlined what sponsors and developers, respectively, expect to learn from evaluations of systems, with modularity and extensibility identified as the most problematic features to assess. Doris Albisser (Union Bank of Switzerland) concluded the session with a prospective customer's view, distinguishing linguistic, technical, organizational aspects and not forgetting the corporate situation of the MT supplier (e.g. its financial stability, resources for system development, and the level of customer support).

The final day of the conference comprised a further two panel sessions and the inaugural meeting of the International Association for Machine Translation (reported elsewhere in this issue.) The first of the panels was devoted to future developments in research: brief descriptions were given by Peter Brown of the statistics-based approach at IBM, by Akira Kurematsu of the speech translation research at ATR, by Hozumi Tanaka of the multilingual CICC project, by Jaime Carbonell of the knowledge-based MT research at Carnegie-Mellon University, and by Masaru Tomita of his plans for training Japanese students in efficient use of existing MT systems. The last panel was entitled 'applications of MT technologies'. Rod Johnson (IDSIA) and Lori Levin (Carnegie-Mellon University) both advocated the use of MT technology in the development of systems for computer-assisted language learning; Yorick Wilks described work on automatic extraction of lexical data from dictionaries; Richard Kettridge (Montreal University) described research on the generation of multilingual texts directly from databanks, i.e. mapping from structured data onto well-defined sublanguages; Hiroshi Uchida (Fujitsu) described the EDR project; Steven Weinstein (Reuters) described the Topic Identification System, which 'skims' texts for 'interesting' news, with the possibility in the future of generating stories in many languages and

thus providing translations of summarized texts; Ian Richardson (IBM Microsoft) outlined the prospects of integrating MT and grammar checkers for authoring-translating systems; and finally Pierre Isabelle spoke about work in Canada on a translator's workstation.

As a diversion from these presentations, the conference provided an opportunity for the delegates to view the latest wares of the MT manufacturers and the latest developments of the research groups. All the well-known names were present: Catena Resource, CICC, ECS, EDR, Eurotra, Fujitsu, Hitachi, Linguistic Products, Logos, Mitsubishi, NEC, New Mexico State University, NTT, Oki, Sharp, Siemens, Systran, Toshiba, TTI (Tovna); and some less well known: Concern Data (with a system for translating into Russian), Intergraph (which has taken over and re-packaged the Weidner MacroCAT system), and Socatra (demonstrating its XLT translation service).

Anyone who was not at the conference can purchase the proceedings from the organizers at US\$40.00 per copy. Orders with a check should be mailed to: Barbara Moore, Center for Machine Translation, 109 Smith Hall, 5000 Forbes Avenue, Pittsburgh, PA 15217, USA.

The fourth MT Summit conference will take place in 1993 in Japan. Plans have already begun and this newsletter will report on progress from time to time.

Evaluators' Forum, Les Rasses, Switzerland, April 1991.

Margaret King

In April of 1991, the International Working Group on Evaluation of Machine Translation Systems (a small working party formed during the Austin, Texas, conference on Theoretical and Methodological Issues in Machine Translation in 1990) organized a four day "Evaluators' Forum" to discuss informally evaluation techniques.

The Forum attracted some 40 participants, representing MT users as well as academic and industrial developers of not yet commercialized MT systems, who came from Japan and from North America as well as from Europe. With a single exception, vendors of already commercialized systems were, sadly, conspicuous by their absence.

The primary purpose of the event was to provide people involved in evaluation with a forum for the exchange of experience and of views on effective evaluation techniques. The discussion was therefore structured around critical presentations of particular methodologies or techniques. On the basis of short contributions submitted before the meeting by the participants, this gave the following main themes:

1. Aspects of overall acceptability of an MT system
2. Reports on some actual evaluations
3. Types of test material
4. Intelligibility, fidelity and other aspects of adequacy of translations
5. Approaches to error analysis
6. Test suites and standard test approaches.

A role-playing exercise served as a way to sum up the discussions: groups of participants tried to determine what the best evaluation strategy would be in a variety of different environments.

Most of the participants seemed to feel that the meeting had been both enjoyable and fruitful, and a second meeting was tentatively scheduled for late January of 1993.

Most of the participants also agreed to supply fuller versions of their contributions. These are currently being collated, and will appear as a technical report in the near future.

Anybody interested either in obtaining a copy of this report or in receiving information on the second meeting when it becomes available should contact: Kirsten Falkedal, ISSCO, 54 route des Acacias, CH-1227 Carouge, Geneva, Switzerland.

1991 Nordic Computational Linguistics Conference focuses on MT

Joseph Pentheroudakis

The 1991 Nordic Computational Linguistics Conference was held on the spectacular campus of the University of Bergen in Norway from the 28th to the 30th of November 1991. This gathering regularly attracts linguists and researchers from Iceland, Denmark, Norway, Sweden and Finland - and, in 1991, from as far away as the United States. This year the meeting was organized jointly by the Norwegian Center for Research in the Humanities (NAVF) at the University of Bergen, and by the University's Linguistics and Computer Science Departments.

Bente Maegaard, Director of the Center for Language Technology at the University of Copenhagen, opened the conference with an overview of the state of the art in the field and an encouraging discussion of the applications that can flow from research in the area. Machine translation was, of course, a distinguished application; in fact, Bente Maegaard was in charge of the Eurotra team at the University of Copenhagen.

Additionally, however, readers of the newsletter will be delighted to know that of the 27 papers presented at the conference, 9 were on MT and related topics, and that two MT systems were demonstrated at the University during the conference. Lars Ahrenberg of Linköping University in Sweden presented a paper on the use of interlingual text representations in machine translation and multilingual text generation in a system implemented at Linköping and used to describe complex devices and their parts in Swedish and English. Björn Beskow of Uppsala University discussed the unification-based transfer in the MULTRA system, an interactive support system for translation and writing under development within the 'Multilingual Support for Translators and Writers' project at Uppsala. The dictionary work for this project, which has Swedish as the source language and English, French, and German as the target languages, was discussed by Annette Østling.

Margrethe H. Møller and Ellen Christoffersen presented their work on the translation of noun phrases from German to Danish, carried out as part of the METAL system's German-Danish project. Joseph Pentheroudakis, of ECS, Inc., in Provo, Utah, discussed some of the development-time limitations in building a machine translation system; ECS's LFG-based Machine Translation Toolkit is installed at the University of Bergen as part of the school's MANTRA project. Gudrun Magnúsdóttir, of Göteborg University, discussed some of the theoretical issues in evaluating machine translation systems, an area in which she has been working for some time now with Margaret King at Geneva. Eva Wikholm (Uppsala University) spoke on translation theory and machine translation.

Roald Skarsten (University of Bergen) and Magnar Brekke, formerly at the University of Bergen and currently at the Norwegian Business School in Bergen, presented the results of the ENTRA project, which involved the computer-assisted translation of several thousand words of petroleum industry texts from English to Norwegian; their results, showing that MT can indeed cut the cost of translation, will be of interest to developers and users alike.

Finally, Helge Dyvik (Department of Linguistics, University of Bergen) presented a paper on linguistics and machine translation, discussing the pivotal role of linguistics in the development of machine translation systems. Helge Dyvik also showed his own unification-based research MT system during the demonstration sessions; the interlingua-based system works between English-Swedish and Swedish-English, and Swedish-Norwegian and Norwegian-Swedish.

The proceedings of the conference will be published. The next Nordic Computational Linguistics Conference will take place in the spring of 1993 at Stockholm University.

MT Workshop in Taiwan, May 1991.

Makoto Nagao

[from JAMT Journal issue no.1, July 1991]

The Taiwan MT Workshop promoted by the National Tsing Hua University took place on 25th-26th May 1991. Scott Bennett, a researcher in MT from the University of Texas, USA, Masaru Tomita of Carnegie-Mellon University and I were invited to give lectures. Keh-Yih Su of the National Tsing Hua University led the collaboration with business to develop the ArchTran English-Chinese MT system. This is the first commercial system in the world. It gives probability scores for grammar rules, etc. with the aim of getting rid of ambiguities. Because of this study, it excels in creating appropriate text. It seems that there are 10 researchers in natural language processing and MT in Taiwan.

MT Vendors Exhibit at the ATA

The 1991 conference of the American Translators Association (ATA) was held in Salt Lake City, from October 16th to the 19th. Thanks to the efforts of Muriel Vasconcellos, Chair of the ATA's Committee on Machine Translation and president of the AMTA, several MT developers and vendors exhibited at the conference: Logos, METAL, ECS, and InterGraph were all there, welcoming streams of interested participants to their booths.

Logos Corporation has ported their system to the SUN SPARCstation, taking advantage of the increased interest in workstation environments. According to Charles Cramer, an interface with Interleaf has been developed, along with interfaces with WordPerfect and FrameMaker. The language pairs that Logos supports are English to French, Spanish, German and Italian (forthcoming), and from German to English, French and Italian.

InterGraph showed their ever-improving DP/Translator, which can now translate CAD/CAM files created under the InterGraph system as well as under third party systems. InterGraph's system includes English to French, German, Spanish, Italian and Portuguese, German to English, and French to English. InterGraph was represented by Tom Reutter and Jacqueline Hannibal, members of the development and customer support staff, and by Ghassan Haddad, manager of the company's translation division.

ECS demonstrated the ECS MT Toolkit, a language-independent programming environment used to implement machine translation systems; the toolkit has been used to develop English-Chinese, English-Japanese, English-Korean and Korean-English, and is currently used by the University of Bergen's MANTRA project to develop a bidirectional English-Norwegian system. ECS also demonstrated its English-Korean bidirectional system as it is integrated in TACCIMS (Theater Automated Command and Control Information Management System), currently installed in Seoul, Korea. ECS was represented by One Soon Her, Dan Higinbotham and Joseph Pentheroudakis.

Finally, W. Scott Bennett of Siemens-Nixdorf and the Linguistics Research Center at the University of Texas, was there to answer questions about the METAL system. In addition to German-English and English-German, METAL offers German-Spanish, Spanish-German, and French-Dutch.

It was a wonderful opportunity for all the vendors that participated to get together and renew old friendships. Exhibiting at the ATA is always a challenge for MT developers: ATA attendees have the most exacting standards and represent one of the best sources for design improvements ideas.

Translating and the Computer 13, November 1991.

John Hutchins

The latest in this series of conferences took place in London on 28th and 29th November 1992, organized by the Aslib Technical Translation Group and the Institute of Translation and Interpreting. Since 1979 the Aslib conferences have provided a forum for translators to learn about developments in MT and computer facilities in the field of translation: multilingual word-

processing, optical character recognition, glossary management and terminology issues. As usual, the conference attracted over 200 participants, mainly from Europe, and contained the familiar mixture of descriptions of research in progress and practical experiences of users. In some previous years the strictly MT content has been slight, but this year a whole day was devoted to MT topics.

The first morning began with an account by Simon Andriessen (INK International) of progress in the 'Language Industries Survey' launched by the Commission of the European Communities in 1987. This comprises essentially a large database of information about private and public enterprises involved in language engineering applications and details of over 1,000 commercially available products. The results are to be published next year as the 'Language Engineering Directory'. The morning continued with two papers on MT: John Hutchins attempted to outline, for those unfamiliar with the field, the basic linguistic complexities and the approaches adopted to overcome them; and Douglas Clarke (Cranfield Institute of Technology) described the advantages of a modular approach to the design of MT systems. The following papers on the first day were devoted to the standards for facilitating the exchange of terminological data (Alan Melby of Brigham Young University), an account of the Eurotermbank at Maastricht (Paul Nekeman), a joint presentation by John Clews (consultant and publisher) and John Parry (Europeanization Computer Products) on the problems of non-English character sets, and two papers by Michael Gavin (Softrans International) and Graham Bason (Lotus Development, Ireland) on the problems associated with the 'localization' of computer software packages, the conversion of commands and menus into other languages and the production of manuals for different environments.

The second day opened with two papers on the evaluation of MT systems. Jacques Durand (Salford University) outlined the basic components (accuracy, robustness, coverage and extensibility) and discussed the development and refinement of criteria for assessing MT systems. Lee Humphreys presented his experience of an evaluation project conducted at Essex University and argued that MT users or potential purchasers should not expect to have to evaluate systems; they should expect systems to perform according to clear specification. The remainder of the second morning was devoted to presentations of MT systems. Nigel Burnford described the facilities available in the Globalink Translation System. François Secheresse described operational experience with Systran, installation and optimization of the system, the role and value of pre-editing, etc. and concluded with an account of recent computational developments (conversion of Systran onto Unix and Sun workstations, rewriting in the C programming language). Thomas Seal described the motives behind ALPNET's change of policy, from producer of machine aids to provider of translation services based on computer technology. It is now acknowledged that most translators do not like the ALPNET/TSS interactive translation facility (Transactive); but they do like to use Autoterm. ALPNET has concluded that systems like TSS are most suitable for large organizations and not individual translators and that the introduction of machine aids requires careful planning and support if the potential benefits are to be achieved.

The final afternoon was devoted to reports on MT research. Makoto Nagao (Kyoto University) gave an excellent survey of current developments in Japan, covering statistics on companies producing systems, the number of systems in daily use, the level of government support for research, the public availability of on-line MT services, and ending with an overview of current research projects. Doug Arnold (Essex University) described the most recent developments in the Eurotra project: the 'pre-industrial' phase ended in 1990, the project is now in a two-year 'transitional programme' with the aim of exploring commercial and research applications. The future is seen in the broader context of the 'linguistic engineering' research programme of the Commission of the European Communities, the development of common tools and resources for multilingual applications and processes, where translation systems represent only one component. The conference ended with a description by Harold Somers (UMIST) of

research on the development of tools to enable monolingual users to compose texts interactively in their own language which will then be automatically translated into another language, with translation quality assured by limiting the domain, controlling input vocabulary and using techniques of example-based MT.

The proceedings of the conference will be published during 1992. The published proceedings of last year's conference "Translating & the Computer 12" are now available (see notice in 'Recent publications' of this issue.)

MMT '91 Beijing

Hirosato Nomura

[from JAMT Journal issue 2, September 1991, p.16]

The International Symposium on Multilingual Translation '91 took place on 19th-21st August in Beijing, People's Republic of China. The venue was Chunichi Seinen Yuko Kaikan (The Young People of China and Japan Friendship Assembly Hall). About 150 people attended including those invited from Japan, Taiwan, America and France. Over 20 people from Japan attended. There were lots of young researchers from China.

MMT'91 is the follow-on from MMT'90 which took place last year at CICC's MT Research Centre in Tokyo. The aim is to advance the project for developing an MT system for translating the languages of neighbouring Asian countries. This time MMT was run in conjunction with the International Symposium for Chinese Information Processing Applications 1991.

Presentations included an introduction to CICC's MT project, progress reports on the current status of the part for which China is responsible, and research on natural language processing in China. The invited speakers from Japan, Taiwan, America and France gave presentations on the current status or problem areas of multilingual machine translation and natural language processing. There were two panel discussions:

- methods of advancing MT R&D from a technical point of view
- progress in MT R&D through control and management

Chinese-English simultaneous interpreters were used, so many Chinese speakers spoke in Chinese and the other speakers used English.

The symposium included exhibitions and demonstrations of MT systems, like CICC's MT systems and the related Chinese system. It is said that at least one MT research project is going on in all the large towns of China. MT research has a long history there and a great deal of interest is shown. Many researchers have studied abroad, they have new computer models like those in Japan and elsewhere and it was clear that research was progressing steadily. The Chinese researchers showed great interest in those from Japan. The symposium provided the opportunity for cordial exchange between the researchers.

PROGRAMMES, POLICIES, SURVEYS

U.S. Commerce Department Meets with MT Vendors and Developers

Joseph E. Pentheroudakis

Dr. Joseph E. Clark, Deputy Director of the National Technical Information Service, United States Department of Commerce, met with representatives of MT vendors and developers during the MT Summit III conference held in Washington, D.C., in July, 1991. The meeting was intended to help identify ways for the US government to support the successful development and use of MT technology.

The meeting was attended by representatives of practically all US developers and vendors: Dragon Systems, ECS, Globalink, InterGraph, Linguistic Products, MCC, METAL, PAHO, TOVNA, Systran, and Xerox were all present. The participants were in agreement that the time is ripe for the government to assist the MT industry in raising its profile and increasing its customer base, which they see as the best avenue towards the development of competitive and more accurate and useful systems.

There were several items on the developers' wish list. Chief among them was the wish to have the government commit to MT technology by encouraging its agencies to use MT for their translation needs. Disseminating information and reports about the successful application of MT would also help, by educating potential users about the usefulness of the technology.

Several groups mentioned that attending conferences and trade shows often represents a substantial financial sacrifice; the government could help by providing direct financial assistance, but also by helping organize low-cost but high-visibility meetings.

The twin issues of standardization and interoperability also came up, both in terms of text encoding standards and in terms of a (loosely) shared linguistic framework. The lack of standards is detrimental to the widespread successful application of the technology. Several national groups are currently working on standardization issues.

Although this is not felt to be a high priority item by developers, Dr. Clark suggested that the government could support research by making large amounts of texts and linguistic data available for testing and system benchmarking. Dr. Charles Wayne, also present at the meeting, is in fact one of the coordinators of the Linguistic Data Consortium, the main objective of which will be the collection and distribution of precisely these kinds of data (see story in this issue of the Newsletter.) Developers, however, generally feel that this sort of 'supply-side' support, if we may be allowed to resurrect the expression, would not be as useful as increased demand-side efforts at raising the industry's profile and carrying out customer education. As Muriel Vasconcellos pointed out, working with customers provides developers with access to an almost unlimited amount of data, more than enough to allow further development and system improvements. Expanding their customer base is the best way for developers to secure both the amount of data needed for further development as well the funds required to support it.

Finally, the issue of a fruitful collaboration between industry and academic research groups also was discussed. The agenda and schedule of academic research groups are, of course, often not quite compatible with those of developers. However, the two can be viewed as complementary, in that developers can often find efficient solutions and applications to models conceived and first tested in academic institutions.

The participants welcomed the government's renewed interest in MT; clearly, the recommendations of the ALPAC report, issued almost two and a half decades ago, have been followed and have borne fruit. What is urgently needed, both for short-term and for long-term benefit, is the kind of government-industry cooperation that will raise the profile of MT, educate the public and help increase the user base for MT systems.

DARPA Linguistic Data Consortium

The US Government is showing continued interest in supporting the development of effective natural language processing technology, including systems for text retrieval, understanding, generation and translation. DARPA, the US Defense Advanced Research Projects Agency, is proposing the establishment of a Linguistic Data Consortium whose purpose will be to develop and make available large amounts of linguistic data on a pre-competitive basis. The data will include large quantities of raw and annotated text and speech, a large lexicon, and a broad coverage grammar of English; the Consortium is also envisioning the eventual acquisition of foreign language materials.

Participation will be broad, and will include many companies and universities, as well as several US government agencies. The data collected will be made available exclusively through the consortium.

Watch for future announcements about this effort, which is under the general direction of Charles L. Wayne at DARPA. An initial request for an expression of interest was issued in July; future announcements will include information concerning general membership and contracts for data production.

JTEC Report on Japanese Machine Translation

The Japanese Technology Evaluation Center (JTEC) of Loyola University, Baltimore, Maryland recently sponsored a study on the state of the art of machine translation in Japan. A panel of American scientists visited 28 Japanese sites, including commercial vendors and users, academic researchers, and government laboratories. Elaine Rich of Microelectronics and Computer Technology Corporation (MCC) reports that the results of the JTEC survey will be ready and publicly available early in 1992 from the National Technical Information Service.

The report covers areas such as the role Japan believes MT will play in assimilating information into Japanese as well as disseminating Japanese information to the rest of the world. Actual statistics of commercial use of MT are included, as well as information concerning the acceptance of MT by users, the status of development of knowledge sources such as dictionaries, and funding for research and development in MT. The report promises to be interesting reading; for information about obtaining a copy, write to: JTEC Report on Machine Translation, National Technical Information Service, US Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161, USA or telephone the NTIS at (703) 487-4650.

JEIDA Survey of MT projects

At the end of 1991 the Japan Electronic Industry Development Association (JEIDA) sent out a 'Questionnaire on the Current Status of Research and Future Trends in Machine Translation and Natural Language Processing'. It is seen as a follow-up of the 1989 JEIDA report entitled "A Japanese view of machine translation in the light of the Recommendations and Considerations Reported by ALPAC, U.S.A." and as part of "a research project on the global status of studies in machine translation and natural language processing..."

The questionnaire, which was to be completed by end of December 1991, has two parts. The first sought the opinions of individual experts on where the current problems of MT are to be found, how soon MT systems are likely to be used by a majority of scientists or business people, and what types of systems are likely to be available in the near future. The second part was a questionnaire about MT and NLP systems and projects, seeking details on the type of project (experimental, commercial), languages involved, system type (batch, interactive), size and nature of dictionaries, source and content of lexical information, the corpus in use, methods of input, morphological analysis, syntactic analysis (grammatical model, parsing method, processing of particular phenomena, disambiguation), semantic analysis (relations, roles, features, use of domain knowledge, etc.), treatment of anaphora, translation method (knowledge based, example based, corpus based, stochastic based), methods of transfer and generation, treatment of non-grammatical input, customization of systems and dictionaries, learning ability, environmental and operational features, evaluations, and future development. When the report appears during 1992 it should provide valuable information for all those involved in MT activities.

European Meeting of Experts Discuss Policy in Language Technology

In November 1991, DG XIII of the Commission of the European Communities (CEC) organized a two-day meeting of over 80 experts to discuss policy in the 'language technology' area. "The panels addressed ways of enhancing and easing human communications by exploiting Information and Language Technologies (spoken and written), and thereby supporting socio-economic developments within the Community and its international competitiveness, while maintaining the diversity of languages and cultures within the Member States... The following Service Areas were emphasized most by the panels: authoring integrated tools, human system interfaces, integrated document management, language engineering, linguistic resources, machine translation, databases, multilingual interfaces and coding, quality assessment, reduced languages, script processing software and system internationalization, speech repositories and terminologies for information exchange. Each primary Service Area was addressed in terms of the needs, problems to be solved, potential solutions and required actions."

The overall conclusion was expressed as the need for a general CEC programme which would cover user and market analysis, technology and information transfer, standards formulation and infrastructure development, building upon existing research and initiatives in the field of 'language technology'.

As far as MT specifically was concerned the major needs identified included:

- the development of system for the translation of patents from "a simplified English source language into, initially one and later all languages used within countries who participate in the European Patent Convention". It could be achieved with the use of PATLING or Machine User Translation Adapted System (MUTAL) - languages with simplified grammar and syntax and with consistent semantics - and the development of patent oriented dictionaries and phrase vocabularies for each technical field.

- methods for measuring quality. It was suggested that quality "may be enhanced through automatic evaluation of performance"

- the development of translator support systems providing the structured storage and extraction of translated texts, "rapid access to standard modules via descriptors or automatic matching, interactive matching of words and parts of text and automatic indexing".

- the need for the "provision of systems for languages with low commercial potential."

- the design of tools for automated acquisition of specific lexica and for performance evaluation.

- the development of "robust morphological analysers and lemmatizers for processing text on a large scale"

- the development of "statistical and probability based techniques to establish measures of similarity between source and archived strings."

- the reuse of lexica and existing linguistic representations.

- an "agreement on specific domain language structures to make them more amenable to machine translation."

Finally, the panel agreed that: "The resource of a language dictionary and its maintenance should be provided on a pan-European basis, thereby lowering development costs for service and product suppliers."

The intention of DG XIII is to produce a 'green book' on language and technology policy in mid-1992 which will serve as the basis for consultation in the Member States. This could lead to a major programme starting early in 1993.

UK Programme in SALT

The UK Department of Trade and Industry (DTI) in conjunction with the Science and Engineering Research Council (SERC) has recently announced the launch of a programme of Speech and Language Technology (SALT) research in the United Kingdom. Its aim is to stimulate research

in this field and to familiarize UK industry with its practical import and potential by encouraging collaborative research between industrial companies and university researchers. There is a particular hope that small and medium sized firms will form an important element in the programme. The funding allocated for projects is £4 million from the DTI and £3 million from SERC for academic participation, matched by £7 million from industry. The programme will extend for three years during which proposals are invited. The intention is to develop UK research capability in areas requiring fundamental research and encourage industry to develop and exploit the results commercially. The main areas to be covered are: spoken language input for human-computer operation, script recognition (analysis of handwritten text for computer input), natural language communication with computer systems (in unconstrained English), improving access to and processing of written documents, treatment of multilingual documentation and "multilingual document production - for use by monolingual users in a wide variety of communicative situations". The programme's work plan is the product of discussions in the UK's SALT Community Club since 1989. It is seen as complementing the European Community's programmes which emphasize international collaboration. The DTI/SERC programme is intended to strengthen the national UK commercial involvement in these technologies.

NEW PRODUCTS

Announcements by Japanese companies

Sanyo Japanese-English/English-Japanese MT system <HEAVEN JE/EJ>

[from JAMT journal issue 2, September 1991 p.14-15]

The Heuristically Engineered Advanced Verbatim ENgine (HEAVEN) is a translation support system suitable for translation-related tasks ranging from machine translation to dictionary reference, editing of texts (pre- and post-editing) and so on. You do not need to use a workstation or dedicated computer; HEAVEN works on a PC (OS/2), so MT is now available at a low cost. Japanese-English and English-Japanese translation is now possible with a single piece of hardware.

Features.

- (1) MT using a PC (OS/2)

Full scale machine translation was previously only possible with a dedicated computer or workstation, but can now be used with a PC

because it works with a PC (using the OS/2 operating system), it is possible to use your existing PC.

you can use your existing texts on MS-DOS files, without any conversion.

you can translate from Japanese to English and English to Japanese with a single piece of hardware.

the translation takes place at high speed - Japanese to English 4000 words/hour, English to Japanese 5000 words/hour.

- (2) An easy to use translation operating environment with GUI The system uses PM OS/2 with a graphics user interface. The translation system can be operated easily using the mouse.

Functions.

batch translation	interactive translation
unknown word search	set up translation mode
dictionary editing	dictionary file management
text editing	

Translation help functions.

These are comprehensive translation help functions - pre-editing functions like word matching, unknown word, split text, no translation etc. and unknown word search.

(3) High speed language processing

Translation.

A unique language for grammar description - flexible grammar rules. The grammar rules are structured using classification of noun meanings or whole sections and the accuracy of word matching analysis has increased.

Translation dictionary.

As well as the basic dictionary, user dictionaries and specialized dictionaries are also available. The basic Japanese-English dictionary contains 57,000 words and the English-Japanese contains 50,000 words. Using the dictionary editing function, you can construct a system compatible with the user operating environment.

Japanese-English/English-Japanese MT system operating environment:

Hardware	32 bit PC (monitor, keyboard, mouse)
Internal memory	8 MB (min)
Hard disk	40 MB (min)
OS	OS/2 version 1.1 or higher

system specifications:

	Japanese-English	English-Japanese
Dictionaries	basic dictionary (57,000 approx.) + user dictionaries	basic dictionary (50,000 approx.) + user dictionaries
Translation speed	4000 words/hr	5000 words/hr
	4 A4 pages (1600 characters/page)	8 A4 pages (600 words/page)

Sharp English-Japanese MT system <DUET Qt>

[from JAMT Journal issue 3, November 1991 p.12]

Features.

1) Proven translation algorithm The Duet Qt utilizes a translation algorithm known as the semantic transfer system

2) 9 comprehensive specialized dictionaries (optional). In addition to the six specialized dictionaries currently available - information processing, electronic engineering, mechanical engineering, economics, chemistry, and general science - Duet Qt now offers three new specialized dictionaries covering medicine, biotechnology and automotive technology.

3) Compact space-saving design. The size of an A4 file, the workstation using the UNIX operating system is the smallest and lightest translation system in use in the business world.

4) English OCR with higher recognition rate (optional) The Duet Qt can be connected to a dedicated English-text OCR with a 99.8% recognition rate (measured by Sharp, varies according to the paper quality and typeface.) By simply moving the mouse, users can indicate the sections of texts they wish to scan, even if the layout is complicated. The ease of document input is thus greatly improved. The OCR can read texts from 6-28 point and recognize 117 typefaces. The spell check can search for typing errors and unknown words automatically.

5) Easy operation with the mouse. This MT software uses the standard Unix X windows system. One-touch menus can be operated with the mouse, making Duet Qt easy to use, even for the beginner.

6) A compact system and high speed translation. In terms of translation quantity, even an expert is limited to a maximum of 2 pages an hour (about 400 words.) Notwithstanding its compact A4 folder size, the Duet Qt can translate at the rate of 12,000 words per hour (about 30 A4 pages.)

7) Compatibility with a variety of word processors and PC's Texts can of course be input using the keyboard or OCR, but English texts created on your word processor can also be input from an MS-DOS formatted 3.5" floppy disk. In addition the finished translation can be copied onto an MS-DOS formatted floppy disk and edited on your PC-based Japanese language WP software.

Pre-editing functions:

1) Specify split translation

She was cute // when she was young.

2) Specify phrases

I saw a beautiful girl with a telescope.

[=by means of a telescope]

I saw <<a beautiful girl with a telescope>>.

3) Specify parts of speech

Help utilities detect errors. [command]

n_Help utilities detect errors. [n_ indicates noun]

4) Specify syntax

He called his son Mike. [asked him to come]

He v5_called his son Mike. [named him Mike]

Hardware:

32-bit CPU (MC68030); 16MB memory; 120MB hard disk (expandable to 200MB); 3.5" floppy disk drive; liquid crystal display monitor; size 320 x 273.5 x 62 mm; weight 3.7 kg; 3 button mechanical mouse; basic software OA/UX 4.0 UNIX system V release 4.0 base X-window.

Software:

Basic dictionary, about 79,000 words; user dictionary, up to 40,000 words per dictionary; learning dictionary available; specialized dictionaries (optional): information processing, electronic engineering, mechanical engineering, economics, chemistry, general science, medicine, biotechnology, automotive technology

Translation method: semantic transfer system.

Speed: 12,000 words/hr, 30 A4 sheets (400 words/page) Mode: interactive/batch translation; standard/large type size; 'dearu'/'desumasu' style

Help functions: part translation, delete, split translation, specify phrases, no translation, specify parts of speech, specify phrases and parts of speech, compare phrases and parts of speech, specify style, dictionary, insert word, spellcheck, help for specifying parts of speech, search and replace

Options - English OCR: 117 fonts recognized, 6 point to 28 point (2 to 10 mm), 300 dpi resolution, 99.8% recognition rate

RESEARCH DEVELOPMENTS

Electronic Dictionary Research Institute

Narumi Chiba
(Head of R&D, EDR)

[from JAMT Journal issue 3, November 1991, p.11]

EDR was established in April 1986 to carry out MITI's electronic dictionary research project 'Pilot research on electronic dictionaries for natural language processing.' The Kiban Gijutsu Kenkyu Sokushin Centre [Centre for the promotion of basic technological research] and eight companies

from the private sector jointly provided funds. The aim is a nine-year R & D project on high level, general purpose electronic dictionaries for a wide range of uses. The research will be used to realize the next generation of natural language processing and intelligent information processing systems, for which general demand has increased in recent years.

Now the EDR project is over half way through and there is a great awareness of the importance of the fundamental technology of electronic dictionaries both in Japan and abroad.

The EDR electronic dictionary is designed for the Japanese and English languages and aims to have 200,000 words in its basic dictionary and 100,000 in its specialized dictionary (the field of information processing); 300,000 in total. It is composed of the following 10 dictionaries.

- Electronic dictionary
 - Word dictionary
 - Basic dictionary
 - Japanese basic dictionary
 - English basic dictionary
 - Specialized dictionary
 - Japanese specialized dictionary
 - English specialized dictionary
 - Concept dictionary
 - Concept system dictionary
 - Concept description dictionary
 - Cooccurrence dictionary
 - Japanese cooccurrence dictionary
 - English cooccurrence dictionary
 - Bilingual dictionary
 - Japanese-English dictionary
 - English-Japanese dictionary

The word dictionary defines headwords based on the meaning of each word and provides morphological, syntactic and semantic information. The concept dictionary expresses the knowledge relating to each concept in a unique form so that the computer can understand it. In other words, the concept system dictionary places each concept in a hierarchy according to superordinate-subordinate relationships. The concept description dictionary expresses the knowledge relating to concepts in a three-part table consisting of two concepts and a relational factor which indicates the relationship between the two concepts. The plan is to create 400,000 concepts and 6,000,000 concept description tables. As many of these concepts can be set up as common characteristics independent of language, they can be applied to an MT system using an interlingua. The cooccurrence dictionary is a collection of phraseological information and is the largest ever of this kind. The bilingual dictionary defines headwords in terms of word-by-word translations.

Dictionary development takes place in two stages. The first stage is to make a vocabulary dictionary in electronic form using vocabulary data made up basically in the same way as for existing dictionaries. In parallel with this, a large number of examples of usage are being gathered and analysed by computer and manually, and syntax trees and concept relational representations are being prepared. As well as this data being collected for the main part of EDR, the data for the cooccurrence and concept dictionaries are being extracted from the same source data. The second stage is the improvement and expansion of the dictionary where vocabulary deficiencies or information errors have come to light as a result of the analysis of the many examples of usage. This will take place automatically or semi-automatically with the aid of computers.

The research into the EDR electronic dictionary is divided among the research laboratories of the companies providing funds and the R&D centre of EDR itself. An effective research structure has been set up using a computer networked dictionary development support system.

The preparation of the dictionary data and the text analysis data requires a lot of manual labour and took place with the cooperation of publishing companies and others. It is a set-up for collecting a great amount of knowledge, developed in cooperation with other leading research establishments and through research partnerships in Japan and abroad.

Centre for the International Cooperation for Computerization

*Yoshinori Teramoto
(Head R&D, CICC MT Research)*

[from JAMT Journal issue 2, September 1991, p.13.]

The CICC MT System Research Laboratory was established within CICC in 1987. It was commissioned by MITI to carry out the ODA project, cooperative research into an MT system for Asian countries. The aim of the project is to break down the language barriers in Asia, so that industrial and technological interchange can be all the more effective. The objective is to develop a multilingual machine translation system which can translate between five Asian languages - Japanese, Chinese, Indonesian, Malay and Thai. This partnership in MT development should produce results such as:

- i) the acceleration of technological and cultural exchange between these neighbouring countries, as the results of the research are applied
- ii) Japan can pass on its advanced information processing technology more effectively and train people in the expertise necessary for computerization
- iii) the establishment of basic computerized technology for word processor development for the other languages, derived from this technological development.

The MT system being developed by CICC uses an intermediate language. This method establishes the semantics and expressions that all the languages have in common and uses this as the core of the translation. This seems the most suitable method for a multilingual MT system:

- i) it allows efficient system development
- ii) it is easy to add languages for translation
- iii) it is possible to develop it for every language. This system uses the field of information processing as a model case, and proves that it is possible to translate the documents in this field at a practical speed and with accuracy. The aim is to achieve a speed of 5000 words per hour when using a standard workstation and an accuracy rate of over 90%, provided the original text is grammatically correct and all the vocabulary is in the system's dictionary. Also, in order to improve the quality of the translation, a system is being set up whereby the translator can edit whenever required.

50,000 words are loaded into the basic dictionary for each language and 25,000 words of specialist vocabulary for the information processing field. The grammar rules are developed based on the analysis and generation of 3000 examples for each language and afterwards improved using a large quantity of texts. The following are also being developed:

an input/output system for inputting text (word processors for every language and a Chinese OCR)

a translation support system for manual or automatic pre- and post-editing

an operational system to integrate and manage these requisite systems

The research receives guidance and advice from Denshi Gijutsu

Sogo Kenkyujo (Institute for Research in Electronic Equipment). It has the cooperation of civil service research establishments in China, Indonesia, Malaysia and Thailand and seven Japanese computer manufacturers (Fujitsu, NEC, Hitachi, Sharp, Toshiba, Oki and Mitsubishi). EDR is cooperating in the development of dictionaries and INTAP is commissioning work related to OSI.

1991 is the fourth year of this six-year research project. During this year we anticipate that we will have reached 40,000 words per language in the basic dictionary, 20,000 per language in the specialist dictionary and 2000 examples for the analysis and creation of grammar rules. With regard to the creation of a multilingual MT system, we should have reached a level whereby

we have strengthened and expanded each of the requisite systems, and integrated them into one MT system. In the future we will need to discuss how to use the MMT system which will come from the project and whether the fundamental results can be used to contribute to computerization in each country.

USERS' VIEWS

A letter to the editor

Claude Bédard (Montréal, Canada)

Long live the new IAMT and its founders! A world association on machine translation, along with a regular publication, has been long overdue - almost scandalously so.

However, I remain concerned that the spirit which I detected at MT Summit III, where the IAMT was launched, should not permeate the future association. My point is that in forums dedicated to MT, like the Summits, the supply side (mostly vendors, along with some major devoted clients) often tends to be over represented, and tends to impose more or less consciously a context of oversimplistic optimism which seems unreal to those who know better, and is simply unfair to those who don't.

This tends to create a climate of intellectual tension, where the "it-works" (the Believers) keep arguing away with the "it-doesn't" (the Detractors). In such circumstances I personally tend to side with the latter - which is absurd indeed since I am really a Believer. I want to MT to work, and not only because we have no other choice anyway considering the potential demand for translation.

Couldn't we create a more relaxed, constructive atmosphere, where the Detractors will simply feel out of fashion, where the Believers will feel free to think realistically, and where we all can address important, insightful issues?

One of these issues is that there are perhaps two markets for MT, i.e. translators and non-translators, which require different design approaches. So far MT systems have really been designed for non-translators (end-users) who are basically faced with the choice between a timely and affordable MT output and an untimely and unaffordable human translation. Professional translators, on the other hand, often confusedly feel that they are not getting from such a machine the right kind of assistance; the machine happens to perform only the trivial aspects of their work (faring even poorly with some time-consuming subtrivialities), leaving them with anything in the translation which is difficult (or subtrivial). Don't they deserve something different from what some of them call an expensive typing help?

Speaking of MT for translators, another important issue then arises: in view of the current limits of MT systems, developers are inclined to think that what translator users then need is more technology. But this implies costlier systems to buy and maintain, which tends to squeeze cost-effectiveness (the bottom line in MT for translators) into a very tight corner indeed. In my view, what we now need as users is a more efficient use of technology; this means on one hand increasing the level of "smartness" among users, and on the other hand tapping ideas from translators in order to develop a technology which will provide them with the right kind of help.

It is high time that the supply side of MT is not left alone defining in its own limited way the terms of reference in this field. Translators are the only ones who know what the task to be performed is all about, yet very seldom do they manage to have anything to say about MT as a tool for them. They have largely failed so far to look into the translation process through the looking-glass of automation.

This situation must change, for everyone's benefit. I therefore urge everyone who has user concerns about MT to join the IAMT with a view to participating in building up a fresh awareness about using and designing MT technology.

RECENT PUBLICATIONS

New Ovum report on NL markets

Engelien, Brigitte & McBryde, Ronnie. *Natural language markets: commercial strategies*. London: Ovum Ltd., 1991. x,292 pp. £725 or US\$ 1,345 (ISBN 0-903969-61-0) [available from Ovum Ltd., 7 Rathbone Street, London W1P 1AF]

Six years after publishing the report by Tim Johnson on the markets for natural language products, Ovum Ltd has published another survey for potential investors. This new report, by Brigitte Engelien and Ronnie McBryde, has attempted to forecast growth in five areas: natural language interfaces to databases, machine translation, contents scanning, text editing tools, and talkwriters. The general conclusion is that after many years of slower progress than expected and despite exaggerated claims and consequent disillusion, there is now genuine promise of realistic and practical solutions. For each area the report lists suppliers, examines and analyses products and charts potential market growth in the France, Germany, the United Kingdom and the United States. The greatest potential is forecast for the 'talkwriter', the word processor with speech input, described as the "dream product... guaranteed to be a commercial success if it can really be made to work."

The predictions for machine translation are rather less optimistic. The authors looked at five systems: Logos, METAL, Smart, Systran and Tovna. Japanese systems were ignored completely, as were also the cheaper microcomputer-based systems such as Globalink TS and PC-Translator. The contention is that MT is effective only with "straightforward, factual text, ideally in a limited subject area or domain..." and "most effective with large quantities of text of the same type." The authors report users with significant productivity gains (particularly shorter translation times), quality improvements (more consistent use of terminology), and organizational benefits (greater control over translation workflow), but there is continuing resistance from professional translators and lack of integration with office systems. Their general conclusion is: "Machine translation works - in the right circumstances, and with the right user commitment. It is still far from perfect..." but "where translator resistance can be overcome - or where translators are employed specifically to work with machine translation and are willing to put in the effort and commitment to make the system work - machine translation can be very effective." Growth in MT is forecast for France, from 28 installations in 1992 to 420 in 2000; for Germany, from 34 (1992) to 600 (2000); for the United Kingdom, from 14 (1992) to 120 (2000); and for the United States, from 24 (1992) to 390 (2000). In the four countries together, revenue is expected to grow from 8 million dollars in 1992 to over 180 million dollars in 2000.

In two appendices, the authors provide information about companies and summarize the experiences of users. Among the companies included are the Carnegie Group (which "in January 1991... announced a multimillion dollar, multi-year machine translation project for an unnamed client"), IBM (with reference to LMT research and use of the Canadian Hansard corpus), Logos (with plans for new systems translating into Italian), Siemens-Nixdorf (mentioning new METAL language pairs under test), and Systran SA (including the Minitel service). Among the users of MT systems are Alpnet, the Canadian Government Translation Bureau (mainly the evaluation of Logos), Perkins Engines (using Weidner MicroCAT with controlled input), and World Bank (evaluation and installation of the Tovna system).

There is no bibliography or index, and the report could be faulted for its selectivity and for uncritical acceptance of some claims by suppliers, but this is a commercial report, informative and succinct, intended for potential investors. The MT community in general will be grateful for independent confirmation of the commercial potential of MT systems from an influential source.

Market Report on MT by Frost & Sullivan

In a press release (November 1990) for the Frost & Sullivan report Machine Translation: a technology impact report (#T041) it is claimed that MT "has languished in the U.S. for reasons of marketing rather than technology," and that, although "artificial intelligence techniques will give MT a boost in the mid-90's, vendors would do well to concentrate on better educating users on how to apply existing capabilities". In the US, it asserts, the "MT industry spends all its money on product development, leaving little for marketing" and contrasts an allegedly more favourable situation in Europe and Japan. "The market is still quite small, with MT accounting for only about \$30 million of a world translation market estimated to be in the \$10-20 billion range." It comments that practical use of MT means restriction to a specific domain or control of input text. "Frost & Sullivan believes that high-quality MT software will migrate to PC-based systems, causing a surge in usage. As the market becomes more realistic about MT's abilities and limitations, there will be more use on interactive workstations where human and machine translators can assist each other." The price of the report is \$1,200, available from: Customer Service, Frost & Sullivan, Inc., 106 Fulton Street, New York, NY 10038, USA. (tel: 212-233-1080). In Europe: Frost & Sullivan Ltd., Sullivan House, 4 Grosvenor Gardens, London SW1W 0DH. (tel: 71-730-3438)

Survey of MT from SCS

SCS-Studie Maschinelle Übersetzung: Grundlagen, Stand und Perspektiven. SCS Informationstechnik GmbH Fachbereich Wissensbasierte Systeme im Auftrag BMFT. Hamburg, 1990. 296pp.

This substantial study of the present situation in MT systems and research was commissioned from SCS Informationstechnik GmbH (Cap Gemini Sogeti) by the Federal German Ministry for Research and Technology (BMFT). It covers the field worldwide, based on a survey of the documentation and interviews with experts. It begins by comparing human and machine translation, outlining the significance of MT for the European Communities, for Eastern Europe, for Japan and for Germany, highlighting the technological applicability of MT research, and describing the potential impact on translation services. There follows a survey of the actual achievement in MT in the form of brief well-balanced descriptions of major systems in Europe (METAL, Systran, DLT, Rosetta), Japan (ATLAS, AS-TRANSAC, PIVOT, PENSEE), and the United States (ALPS, Logos), and of major research efforts (Eurotra, CICC, ATR, EDR, KBMT, LMT). Next come chapters on the typology of systems, the linguistic foundations (with excellent summaries of current linguistic and AI theories), and the computational aspects.

The report's conclusions (pp.145-158) emphasize the necessity for large scale electronic dictionaries, particularly for German, if the German expertise in MT research at Stuttgart, Berlin and Saarbrücken (thanks largely to involvement in Eurotra) is to be built upon. The report stresses the importance of fostering university-industry links in Germany, the promotion of international collaboration, and the exploitation of the favourable position of multilingual Europe (compared with the largely monolingual United States and Japan). Most strongly advocated is the future application of MT in systems for searching foreign-language databases and translating the documents retrieved, and thus giving German users access to information otherwise neglected for linguistic or financial reasons. As the most ambitious goal the report sees systems for the automatic

translation of spoken language (eventually on pocket or portable devices.) Before then, it forecasts the general public availability of systems producing 'informative' translations of telex and email messages.

The appendices include a useful glossary of terminology (pp. 169-188), single page overviews of most systems and computer-based dictionaries (pp. 189-246), and a good bibliography (pp. 247-289).

NEW BOOKS

Two on KBMT from CMU

Goodman, Kenneth & Nirenburg, Sergei (eds): *The KBMT project: a case study in knowledge-based machine translation*. San Mateo, Ca.: Morgan Kaufmann Publishers, 1991. xvi,331 pp. \$ 34.95 (ISBN. 1-55860-129-5)

Nirenburg, Sergei, Carbonell, Jaime, Tomita, Masaru, and Goodman, Kenneth: *Machine translation: a knowledge-based approach*. San Mateo, Ca.: Morgan Kaufmann Publishers, 1992. xiv,258 pp. \$ 39.95 (ISBN. 1-55860-128-7)

In the last quarter of 1991, Morgan Kaufmann has published two volumes on knowledge-based machine translation by members of the Center for Machine Translation at Carnegie-Mellon University, Pittsburgh, US. The first edited by Kenneth Goodman and Sergei Nirenburg brings together under one cover the articles previously published in a double issue of volume 4 of *Machine Translation*, which describe in detail the research system KBMT-89. The second is a monograph by four authors (Sergei Nirenburg, Jaime Carbonell, Masaru Tomita and Kenneth Goodman) which is undoubtedly the most substantial publication on the knowledge-based approach in general and on its appropriateness for domain-specific and speech translation in particular. The chapters are headed: MT in a nutshell (p.1-39), Treatment of meaning in MT systems (p.41-63), The concept of interlingua (p.65-93), Lexicography and knowledge acquisition (p.95-116), Source language analysis (p.117-134), Target language generation (p.135-164), Speech-to-speech translation (p.165-190), Machine-aided translation (p.191-204), and The future of machine translation (p.205-213). It ends with a 'KBMT glossary' (p.215-224) and a bibliography (p.225-248). As the authors state, this is "the first monograph devoted exclusively to theoretical and methodological issues in knowledge-based MT" and it will doubtless be essential reading for all MT researchers.

German introduction to MT

Schwanke, Martina: *Maschinelle Übersetzung: ein Überblick über Theorie und Praxis*. Berlin: Springer-Vlg., 1991. 489pp. (ISBN 3-540-54186-1). DM.35,00

In this introduction to MT Martina Schwanke (Kiel University, Germany) devotes short well-written chapters to MT and its relations with translation studies (p.11-45), defining terms and concepts in the MT field (p.47-67), and a history of MT (p.69-84). This chapter is followed by an extract from the first published MT article, by Oswald and Fletcher (1951) on German syntax (p.85-90), and a description of TAUM-Meteo in 1979 (p.91-95). Chapter 4 consists of detailed descriptions of LOGOS, SUSY, METAL, and SYSTRAN (p.97-195), and chapter 5 is a brief outline of computer aids for the translator (p.197-207). The book is completed by a comprehensive directory (System-Dokumentation) of MT systems and term banks (p.209-394) and by an impressive bibliography (p.395-489). There is surprisingly no index.

For German readers this represents probably the most accessible up-to-date reference source for MT in general. It will be valuable in particular for the directory of MT systems and for its substantial bibliography.

OTHER RECENT PUBLICATIONS

Translating and the computer 12: applying technology to the translation process. Edited by Catriona Picken. Proceedings of a conference jointly sponsored by Aslib, The Association for Information Management, Aslib Technical Translation Group, The Institute of Translating and Interpreting, 8-9 November 1990, CBI Conference Centre, London. London: Aslib, 1991.

The volume includes on p.65-82: Language technology products in the European market (Brigitte Engelién) - p.83-109: Machine translation of natural language: the TOVNA MTS solution (Ami Segal) - p.149-158: Translation quality - how can we tell it's good enough? (Peter Kahl)

Machine Translation Summit III: Proceedings, program, contributed papers, panel statements. July 1-4, 1991, Washington, D.C. (Host institution: Center for Machine Translation, Carnegie Mellon University, Pittsburgh, Pa 15213). 156 pp. US\$40.00. [available from: Barbara Moore, Center for Machine Translation, 109 Smith Hall, 5000 Forbes Avenue, Pittsburgh, PA 15217, USA.]

Ian Tresman: *Multilingual PC directory: a guide to multilingual and foreign language products for IBM PCs and compatibles.* Borehamwood, Herts. (UK): Knowledge Computing Ltd., 1991. 256 pp. £19.95 (ISBN. 1-873091-01-x) [available from: Sesame Computer Projects, 8 Avenue Road, Harrogate, N.Yorks HG2 7PG, U.K.]

This is an invaluable guide to language software for IBM PCs, it includes brief notes on M(A)T systems (Globalink TS, Language Assistant, LinguaWrite, Smart Translator, Toltran, Translate, Translator); addresses of companies, agencies and suppliers throughout the world.

Proceedings of International Workshop on Electronic Dictionaries, November 8-9, 1990, Oiso, Kanagawa, Japan. Tokyo: Japan Electronic Dictionary Research Institute Ltd., 1991. (Technical Report TR-031). [available from: Japan Electronic Dictionary Research Institute Ltd., Mita Kokusai Bldg., 4-28 Mita 1-chome, Minato-ku, Tokyo 108, Japan.]

Other reports on the EDR project available are: TR-024 An overview of the EDR electronic dictionaries; TR-025 Japanese word dictionary; TR-026 English word dictionary; TR-027 Concept dictionary; TR-029 Bilingual dictionary; and in course of revision: TR-028 Cooccurrence dictionary; TR-030 Support system for dictionary development.

Studies in Machine Translation and Natural Language Processing. Edited by C. Copeland, J Durand, S. Krauwer, B. Maegaard. Bruxelles/Luxembourg, Commission of the European Communities, 1991. (ISSN 1017-6586)

volume 1: The Eurotra linguistic specifications. p.7-14: An introduction to the Eurotra programme (Bente Maegaard, Sergei Perschke) - p.15-123: Linguistics for machine translation: the Eurotra specifications (Valerio Allegranza, Paul Bennett, Jacques Durand, Frank van Eynde, Lee Humphreys, Paul Schmidt, Erich Steiner) - p.125-160: Dictionaries in Eurotra (Pius ten Hacken, Heinz Dieter Maas, Bente Maegaard) - p.161-163: Terminology in Eurotra (Jennifer Pearson, Dorothy Kenny)

volume 2: The Eurotra formal specifications. p.7-40: Description of the Eurotra framework (Annelise Bech) - p.41-111: The Eurotra user language (Giovanni Malnati, Patrizia Pagio) - p.113-133: The Eurotra software environment: a broad overview (Roberto Cencioni)

Journals:

JAMT Journal (items marked * are reproduced in translation elsewhere in this issue)

Issue 1, 25 July 1991: Messages to the JAMT first ordinary general assembly (Makoto Nagao and Nobuhiro Miyake of MITI) -- Conference reports: *MT evaluators' forum (Switzerland); *MT workshop (Taiwan); *International Conference on Current Issues in Computational Linguistics (Penang) -- Report on the MT SUMMIT III (Washington, D.C.); JAMT regulations; *JAMT organization

Issue 2, 25 September 1991: Report of the JAMT first ordinary general meeting -- Speech by Makoto Nagao at the JAMT meeting -- Introduction to MT history -- Report of activities on the MT frontiers -- Report of papers given at the MT SUMMIT III (Washington, D.C.) -- *The CICC MT project -- *The new MT system HEAVEN JE/EJ -- *Announcement MT World '92 to be held by JAMT, with exhibitions & seminars, March 1992 -- *Conference guide -- Notices from the Secretariat of JAMT activities -- Questionnaire (Voices from JAMT members)

Issue 3, 25 November 1991: Cat and a bad wife (Susumu Donomae) -- Introduction to MT technology -- User guide -- From the battlefield of translation [with English translation] (Tetsuo Sagawa) -- The F1 machine and MT -- Article comparing MT development with the development of new car models -- *Introduction to EDR activity -- *Introduction to a new MT system, DUET Qt -- *Conference guide -- *Announcement on MT World '92 -- Voices from JAMT members -- Notices from the Secretariat on JAMT activities -- List of JAMT members (as of November 10, 1991)

La Tribune des Industries de la Langue numero special no.4-5-6 (novembre 91): Ingenierie linguistique: Problématiques 1995.

The proceedings of the colloquium organized for the Salon international des industries de la langue, Paris, 20-21 November 1991; covering the whole range of products, applications and research involving natural language processing with particular emphasis on the French language. Includes contributions on IBM-Europe, GSI-Erli, and SITE's projected Eurolang software package incorporating a component based on GETA-Ariane.

Machine Translation *vol.6 no.1, March 1991* p.1-20: Relational-grammar-based generation in the JETS Japanese-English machine translation system (David E. Johnson and Hideo Watanabe) - p.21-33: Machine translation of non-literary texts: some Canadian experiences (C.C. Gotlieb and L.d'Haenens) - p.35-54 Book reviews (Harold Somers, Natsuko Holden, Patrick Saint-Dizier, and Angela Ralli)

vol.6 no.2, June 1991 (Special issue on Eurotra I) p.61-71: Introduction (Valerio Allegranza, Steven Krauwer and Erich Steiner) - p.73-82: Eurotra: general system design (Bente Maegaard and Sergei Perschke) - p.83-101: The Eurotra MT formalism (Anneliese Bech, Bente Maegaard and Anders Nygaard) - p.103-147: The Eurotra linguistic specifications: an overview (Jacques Durand, Paul Bennett, Valerio Allegranza, Frank van Eynde, Lee Humphreys, Paul Schmidt and Erich Steiner)

vol.6 no.3, September 1991 (Special issue on Eurotra II) p.149-170: Implementing monolingual grammars and transfer components in the Eurotra formalism (Sabine Kirchmeier-Andersen) - p.171-182: A treatment of unbounded dependencies in Eurotra with an experimental coindexation tool (Valerio Allegranza and Anneliese Bech) - p.183-192: Token reference vs. type reference: implications for machine translation (Cornelia Zelinsky-Wibbelt) - p.193-200: Transfer formalisms (Doug Arnold and Louisa Sadler) - p.201-214: An overview of MiMo2 (Gertjen van Noord, Joke Dorrepaal, Pim van der Eijk, Maria Florenza, Herbert Ruessink and Louis des Tombe) - p.215-

228: CAT2 - an experimental Eurotra alternative (Randall Sharp) - p.229-242: Book reviews (Blaise Nkwenti Azeh, George Dunbar, Bill Black).

selected articles from other journals

Applied Computer Translation vol.1 no.1 (Jan-Mar 1991), p.15-27: The adequacy of corpora in machine translation (Mark Sebba) -- vol.1 no.2 (May 1991), p.35-44: Book review [of] John Laffling Towards high-precision MT, based on contrastive textology (Toon Witkam) -- vol.1 no.3 (July-Sept 1991), p.49-59: Declarative evaluation of an MT system (Lorna Balkan, Matthias Jaschke, Lee Humphreys, Siety Meijer and Andy Way) -- vol.1 no.4 (Oct-Dec 1991), p.5-21: Structural transfer and unification formalisms (Louisa Sadler)

Computational Linguistics vol.17, no.3 (September 1991), p.325-326: Erratum to 'A statistical approach to machine translation (Peter Brown, et al.)

Language Industry Monitor no.1 (Jan-Feb 1991), p.4-5: Japan's chances in the great machine translation stakes (Tony Whitecomb). -- no.2 (March-Apr 1991), p.5-6: Low-cost automatic translation: an American dream? [on PC-Translator] -- no.3 (May-June 1991), p.1-2: Arabic language technology; p.5: MT evaluator's forum; p.6: MT research revitalized from the Hague? -- no.4 (Jul-Aug 1991), p.1-3: What do translators want? (Claude Bédard); p.3-4: Eurotra continues; p.6-7: HICATS/JE for RPI's JIRS ['information tracking' by Research Policy Institute, Lund] -- no.5 (Sep-Oct 1991), p.3: TERM-Trans; p.5: Opening up Eurotra (Tony Whitecomb); p.7: Third MT Summit -- no.6 (Nov-Dec 1991), p.1: New NLP report from Ovum.

Language International vol 3, no.4 (July 1991), p.8: Evaluation of machine translation systems (Monique C.Cormier) [report of ISSCO workshop]

Reports from MT research groups:

Linguistics Research Center (University of Texas) *Quarterly Report*, June 1991, p.1-14: METAL: past, present and future (Winfield Scott Bennett)

University of Essex (Department of Languages and Linguistics) *Working Papers in Language Processing* no.16 (Dec 1990): User-oriented evaluation of MT systems (R.Lee Humphreys) -- no.17 (Jan 1991): Structural transfer and unification formalism (Louisa Sadler) -- no.18 (Jan 1991): Structural non-correspondence in translation (Louisa Sadler) -- no.19 (Jan 1991): Assessing a PC-based commercial MT system (Essex MT Evaluation Group) -- no.20 (Feb 1991): Widgets don't wear their names on their sleeves (R.Lee Humphreys) -- no.21 (Mar 1991): Developer-oriented evaluation of MT systems (Andrew Way) -- no.22 (Apr 1991): Operational evaluation of MT (Lee Humphreys, Matthias Jäschke, Andrew Way, Lorna Balkan and Siety Meyer) -- no. 23 (Apr 1991): Quality criteria for MT (Lorna Balkan) -- no.25 (Apr 1991): Declarative evaluation of an MT system: practical experiences (Lorna Balkan, Matthias Jäschke, Lee Humphreys, Siety Meijer and Andy Way) -- no.26 (Jun 1991): A practical developer-oriented evaluation of two MT systems (Andrew Way) -- no.27 (Sep 1991): Automatic test suite generation (Doug Arnold, Dave Moffat, Louisa Sadler and Andrew Way) -- no.28 (Oct 1991): Multi-word units in linguistic theory and in MT (Tim Nicolas) -- no.29 (Oct 1991): A constraint-based approach to translating anaphoric dependencies (Louisa Sadler and Doug Arnold) -- no.30 (Oct 1991): Knowledge for MT from a general translation theory perspective (Lorna Balkan) -- no.31 (Nov 1991): Valency-changing alternations (R.Lee Humphreys) Internal Memoranda, Computational Linguistics & Machine Translation Group -- no.25 (May 1991):

Research proposal: NP-structure and NP-interpretation (Andrew Way, Alex Theofilidis, Michael Grabski, Toni Badia, Jesus Vidal and Heleen Koekstra) -- *no.26 (July 1991)*: Transfer problems (Daniel Brockmann) -- *no.27 (June 1991)*: Report on ICAME (International Computer Archive of Modern English) Open Day, 9 May 1991 (Tim Nicolas) -- *no.28 (July 1991)*: A guide to the Eurotra translation system: seminar conducted at DCU 22-23 July 1991 (Andy Way) -- *no.29 (Sept 1991)*: Eurotra-UK final report (1985-90) and comprehensive bibliography (1983-90) (Doug Arnold and Jeannette Pugh)
