Verbmobil

Translation of Face-to-Face Dialogs

Wolfgang Wahlster

June 1993

Deutsches Forschungszentrum
für
Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern und Saarbrücken is a non-profit organization which was founded in 1988 by the shareholder companies ADV/Orga, AEG, IBM, Insiders, Fraunhofer Gesellschaft, GMD, Krupp-Atlas, Mannesmann-Kienzle, Nixdorf, Philips and Siemens. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct *systems with technical knowledge and common sense* which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- ❏ Intelligent Engineering Systems
- ❏ Intelligent User Interfaces
- ❏ Intelligent Communication Networks
- ❏ Intelligent Cooperative Systems.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Friedrich J. Wendl
Director

Verbmobil Translation of Face-to-Face Dialogs

Wolfgang Wahlster

Table of Contents

Verbmobil
Translation of Face-To-Face Dialogs

Wolfgang Wahlster

German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3
D-66123 Saarbrücken 11, Germany

phone: +49 681 302 5252 or 2363
fax: +49 681 302 5341
Internet: wahlster@dfki.uni-sb.de

Abstract

Verbmobil is a long-term project on the translation of spontaneous language in negotiation dialogs. We describe the goals of the project, the chosen discourse domains and the initial project schedule. We discuss some of the distinguishing features of Verbmobil and introduce the notion of translation on demand and variable depth of processing in speech translation. Finally, the role of anytime modules for efficient dialog translation in close to real time is described.

The long-term vision behind the project Verbmobil is a portable translation device that you can carry to a meeting with speakers of other foreign languages and it will translate what you say for them.
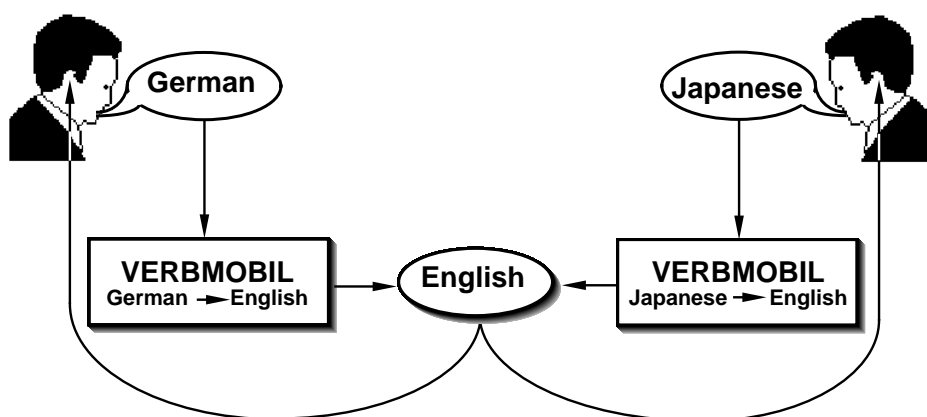

Fig. 1: English as the Common Dialog Language in Verbmobil

This very ambitious scientific goal will be persued in a series of well-defined project phases. The first versions of Verbmobil will provide translation on demand for the two participants who have a passive knowledge of English but of which neither is a fluent speaker. We assume that most of the dialog will be conducted in English as a common dialog language. This is a realistic assumption for most international technical or business discussions. But for uncommon words or phrases, complex constructions and critical segments of the negotiation dialog the participants may want to switch back to their native language. This means that they need translation help and therefore turn to their Verbmobil devices.

In the course of the conversation each dialog partner can activate his version of Verbmobil (eg. German-to-English or Japanese-to-English translation) and signal that he is now speaking in his native language (eg. German or Japanese), and that what he says should be translated into English (see Fig. 1).

This means that there are three input modes for Verbmobil:

1) Both dialog participants speak English with a German or Japanese accent. In this case, no translation is necessary, but Verbmobil has to follow the conversation and extract context information for subsequent translation tasks. This is an extremely difficult problem, since the input can be ill-formed in many ways, so that various phonetic and grammatical constraints have to be relaxed in order to cope with the foreign accent and unusual constructions. Often Verbmobil will extract only a very shallow discourse model using keyword spotting or other partial analysis techniques.

2) In the course of an utterance, a participant switches from English as the common dialog language back to German or Japanese as his native language. In this case, Verbmobil must generate a translation that fits with the context of the English sentence fragment. For example, if a German participant says "Let's meet again in June ↓außer am Pfingstmontag↑" Verbmobil should produce "except on Whit Monday" to complete the English fragment correctly (the arrows ↓ and ↑ indicate that the speaker has signalled the code switching to Verbmobil).

3) The participant speaks in his own language and Verbmobil will translate his utterance into English. In this case, Verbmobil must generate an appropriate approximation of the communicative intent of the input, in close to real time. In many situations, Verbmobil will be able to find translations that preserve most but not necessary all

of the content of the original, since translation is inescapably a matter of compromise.

Clarification subdialogs play an important role in the conversational setting discussed above, since the dialog partners are no fluent speakers of English and Verbmobil is an imperfect understander and translater. In the Verbmobil project, two types of clarification subdialogs are studied (see Fig. 2):

1) Clarification subdialogs between the participants are conducted in English. There are two variants of this type of subdialog: both dialog partners use English or Verbmobil translates their utterances from their native language into English.

2) Clarification subdialogs between Verbmobil and one participant are conducted in the native language of the respective dialog partner.
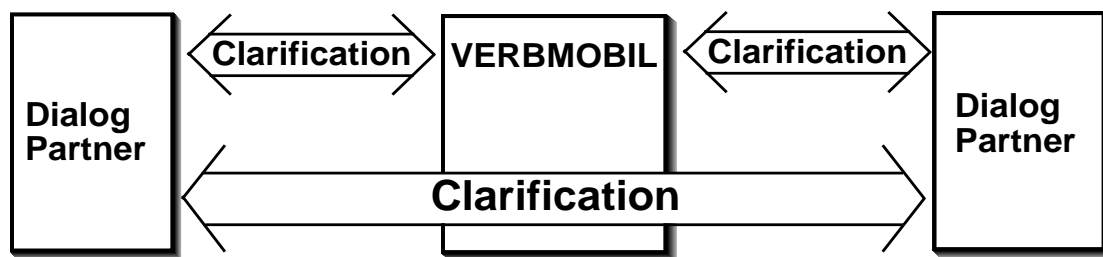

Fig. 2: Two Types of Clarification Subdialogs

The Project Goals

There are four distinguishing features of the Verbmobil approach:

- ❏ speaker-adaptive recognition of spontaneous speech negotiation dialogs in face-to-face situations
- ❏ portable translation device that can be tailored to the individual user and to specified application domains
- ❏ three language scenario (English, German, Japanese) with English as a dialog language, ensuring system transparency and user acceptance.

In contrast to previous projects on speech translation (cf. [2], [5]) Verbmobil does not deal with telephone conversations but with face-to-face dialogs in a small meeting room. In face-to-face dialog translation we can exploit the fact that information passes between the participants

not only on the linguistic channel but also on various nonverbal and paralinguistic channels. The hearer can merge information from the translation with information from gestural motions of the hands, fingers, head and eyes, eyeblinks, eyebrows movements, change of body posture and orientation. The research program includes some empirical investigations of translation and interpreting as done by humans in similar situations.

Verbmobil does not deal with read speech input, but with incrementally produced spontaneous dialog contributions. Such utterances are rarely well-formed, since speakers make errors and correct them. Verbmobil has to deal with false starts, aborted phrases, speech repairs, hesitations, interjections, self-correction phrases and many other characteristic features of spontaneous speech (see Fig. 3).

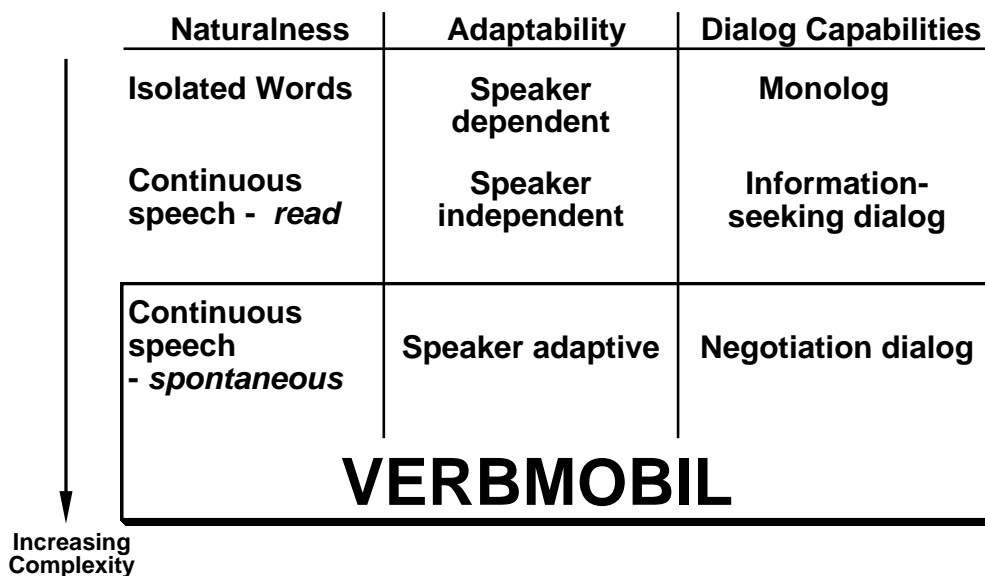| Naturalness	Adaptability	Dialog Capabilities
Isolated Words	Speaker dependent	Monolog
Continuous speech - *read*	Speaker independent	Information-seeking dialog
Continuous speech - *spontaneous*	Speaker adaptive	Negotiation dialog
VERBMOBIL		

Increasing Complexity

Fig. 3: Challenges of Language Technology

In the discourse situation studied for the initial demonstrator the dialog partners discuss a possible date for their next meeting using a calendar in front of them. After the development of the initial demonstrator, the domain of discourse will be extended considerably for the first research prototype. Two negotiation tasks will be considered for the research prototype (see Fig. 4).

Note that the appointment scheduling task is a subtask of both scenarios considered for the research prototype. The domains chosen deal with linguistically ordinary language, so that the linguistic knowledge sources can simply be extended when the domain is scaled up. In all

conversational settings studied in the Verbmobil project the subject matter is limited and the aims of the dialog partners are known in advance. We take it that both dialog partners come to a meeting in a spirit of cooperation and that they are highly motivated to reach a successful conclusion.

Domain Restrictions

❏ Limit dictionary size (1500 words for first demonstrator)

❏ Limit ontology

❏ Facilitate massive and decentralized data collection

❏ Make negotiation dialog possible

❏ Facilitate scale-up

Discourse Situation for Demonstrator (after 2 years)

The dialog partners discuss about a possible date for their next meeting using a calendar in front of them.

Domains for Research Prototype (after 4 years)

❏ Two businessmen plan a joint business trip using various schedules.
❏ A customer negotiates with a sales agent about ordering an item shown in a catalog.
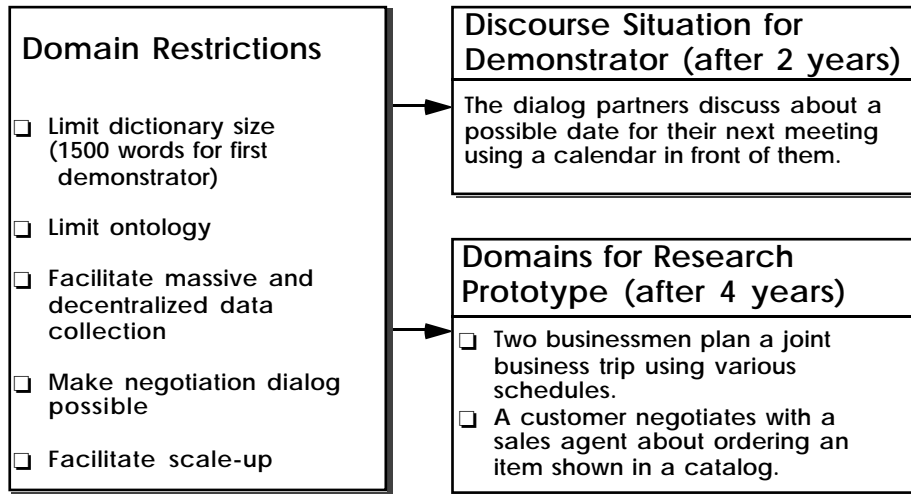
Fig. 4: Discourse Domains for Verbmobil.

Verbmobil channels energy into key areas of language technology and integrates major subfields of advanced information technology like natural language processing, speech recognition and synthesis, machine translation, dialog and knowledge processing (cf. Fig. 5).
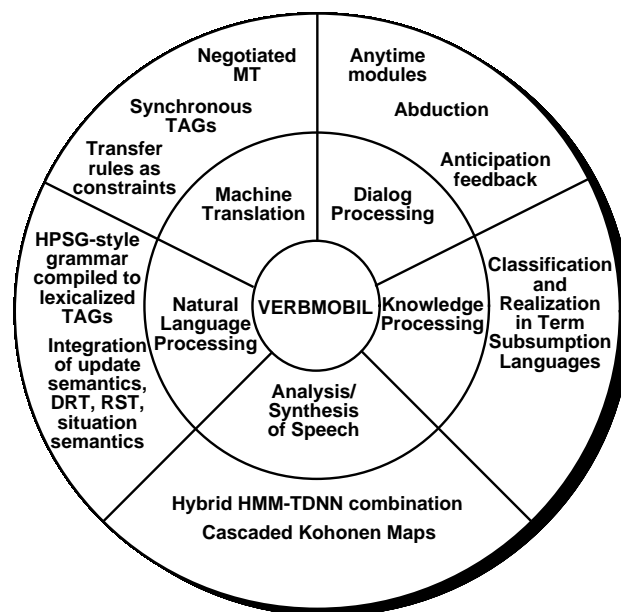
Fig. 5: Integrating Major Subfields of Language Technology

5

Since there in no doubt, that the fact that language is always situated is very important for translation and that a proper translation almost always depends on context, Verbmobil must integrate research on translation with work on dialog processing as well as knowledge representation and reasoning.

Verbmobil is an interdisciplinary attempt to build a face-to-face translation system on the basis of current theories that leading researchers in artificial intelligence, computational linguistics, speech processing, neuro-computing and translation science would subscribe to. The Verbmobil consortium believes that the scientific foundation of dialog translation technology should never be compromised in the interests of achieving some functionality or speed-ups in the short run by ad hoc techniques, that cannot be generalized and scaled-up.

Anytime Modules for Face-to-Face Dialog Translation

Obviously, there is a tradeoff between run-time and quality of results in systems for face-to-face dialog translation. Verbmobil's analysis should not be deeper than necessary, its translation should be as shallow as possible, and its generation process should start as soon as possible.
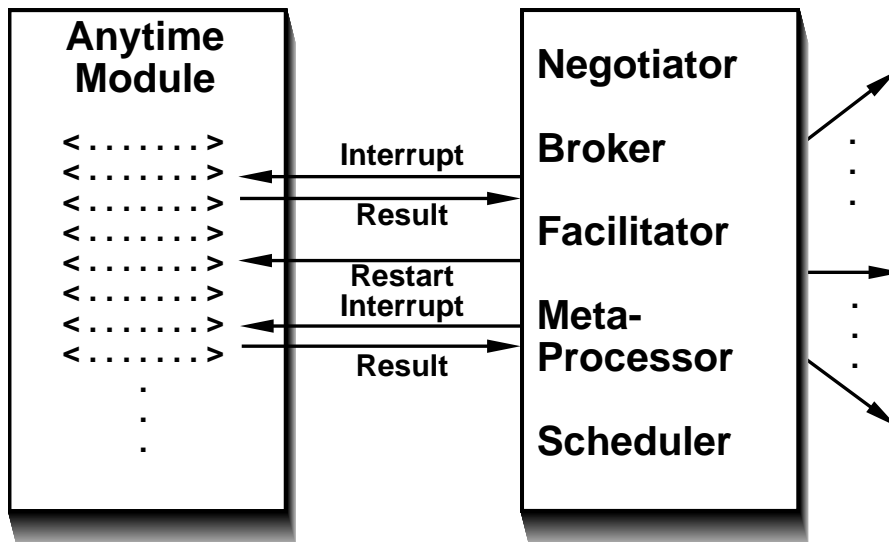
Fig. 6: Anytime Modules as Coroutines

This means that the major components of the system must work in an incremental mode allowing the immediate processing of parts of a stepwise provided input. These modules will be realized as anytime modules for the sake of resource-bounded processing of discourse.

Anytime modules are modules whose quality of results improve gradually as computation time increases. They yield imperfect but not useless results if interrupted before completion. If an anytime module is restarted, it can improve what is has generated so far.

For Verbmobil anytime modules are needed on various levels of granularity, e.g.

- ❐ speech analysis, parsing, transfer, generation
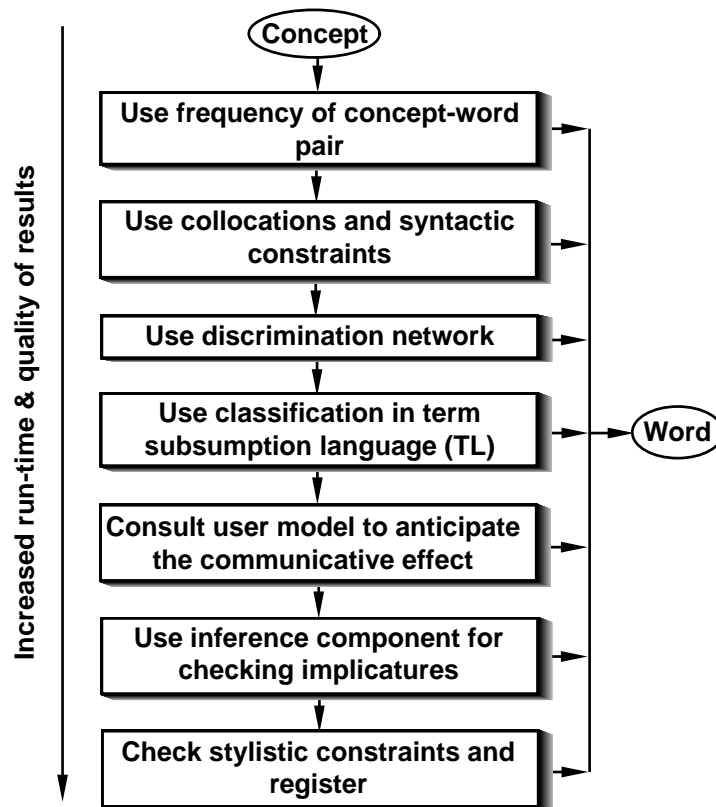- ❐ pronoun resolution, focus detection, lexical choice

Fig. 7: Lexical Choice as an Anytime Module

All Verbmobil modules integrate a wide spectrum of layered methods: from simple and low cost to complex and expensive techniques. This can be illustrated by the problem of lexical choice. If lexical choice is implemented as an anytime module, the quality of the results can be measured in terms of the precision of communicating the intended concept in a given situation. The concept-to-word mapping can be achieved by a wide spectrum of techniques from very fast methods using the frequency of concept-word pairs to very elaborate methods like checking possible communicative effects and implicatures.

7

The concept of anytime modules is tightly connected to the idea of variable depth of processing in a speech translation system. Verbmobil will use a multi-layered semantic respresentation language, that allows for all kinds of underspecification in the surface-oriented layers. In many cases, ambiguous quantifier scope or PP attachment in the source language need not be resolved before being translated, since a corresponding ambiguity can be captured in the target language. This leads to the new problem of language generation from disjunctive semantic structures.

It is important that each layer of the semantic representation language comes with a specialized inference component, so that even on the level of surface-oriented representations simple inferences can be drawn. While these inferences may be based on primitive rewriting techniques, the inference engine on the more elaborate levels of meaning represenation may be a full theorem prover.
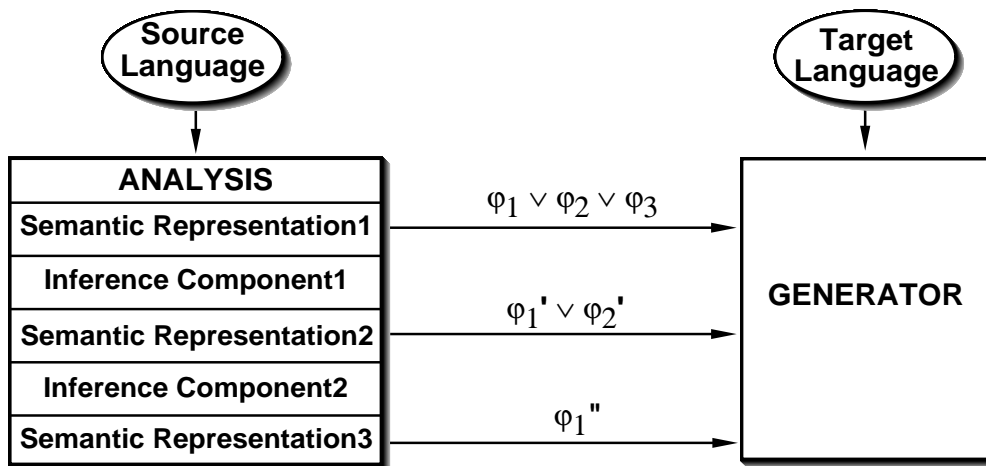

The diagram shows: Source Language → ANALYSIS block containing Semantic Representation1, Inference Component1, Semantic Representation2, Inference Component2, Semantic Representation3. Arrows to GENERATOR: $\varphi_1 \vee \varphi_2 \vee \varphi_3$, $\varphi_1' \vee \varphi_2'$, φ_1''. GENERATOR → Target Language.

Fig. 8: The Notion of Variable Depth of Processing

The Project Structure

The Verbmobil project is funded by the German Ministry for Research and Technology (BMFT) and an industrial consortium. For the first four years of the project the BMFT funding amounts to 60 Million Deutschmarks.

The BMFT commissioned two feasibility studies on the goals of Verbmobil: one from a consortium of German industrial and academic research groups (see [3]) and another from the Center for the Study of

Language and Information (CSLI) in the US (see [1]). Based on the positive recommendations of the two independent studies a detailed project plan and schedule was prepared (see [4]), that formed the basis of a call for proposals in July 1992. An international advisory and review board was appointed by the BMFT consisting of 10 well-known experts in speech, language and translation technology. The scientific review of all submitted proposals was finished at the end of January 1993. The main phase of the project is starting in May 1993.

The project is planned for 8 to 10 years and the first phase of 4 years is structured by 2 major milestones: a demonstrator after 2 years and a research prototype after 4 years (see Fig. 9). The central project coordination task and the implementation of the demonstrator and research prototype will be carried by the German Research Center for AI (DFKI).
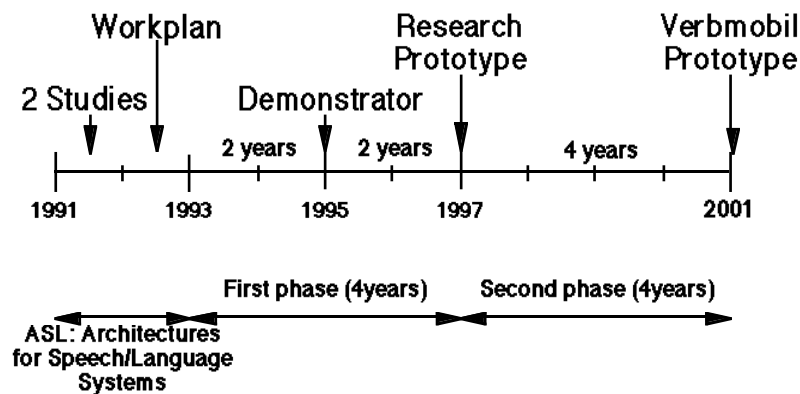
Fig. 9: The Project Schedule for Verbmobil

The success of such an ambitious translation project obviously depends on international cooperation. It is planned to have an intensive collaboration with the ATR Interpreting Telecommunications Research Laboratories in Kyoto. In March 1993, this well-known Japanese center for speech translation research started a new project that will end in March 2000. Like Verbmobil this project deals with the translation of spontaneous dialog language. The funding amounts to 16 billion yen. Data collection, speech modules and linguistic knowledge sources for the Japanese language are the major areas of the planned collaboration. For work packages concerning the English language, cooperations have been prepared with three US research groups: Carnegie Mellon University, CSLI at Stanford University and the International Computer Science Institute (ICSI) at Berkeley.

References:

[1] Kay, M., Gawron, J.M., Norvig, P.: Verbmobil: A Translation System for Face-to-Face Dialog. BMFT Study, CSLI, Stanford Univ., August 1991.

[2] Morimoto, T., Shikano, K., Iida, H., Kurematsu, A. : Integration of Speech Recognition and Language Processing in the Spoken Language Translation System SL-TRANS. In: Proc. of the Intern. Conference on Speech and Language Processing, 1990, p. 921 - 928.

[3] Verbmobil-Consortium: A Portable Translation Device. BMFT Study, Siemens, Munich, August 1991 (in German).

[4] Wahlster, W., Engelkamp, J. (eds.): Scientific Goals and Networks of Work Packages of the Verbmobil Project. BMFT Study, DFKI, Saarbrücken, April 1992 (in German).

[5] Waibel, A., Jain, A.N., McNair, A..E., Saito, H., Hauptmann, A., Tebelskis, J. (1991): JANUS - A Speech-to-Speech Translation System Using Connectionist and Symbolic Processing Strategies. In: Proc. of the 1991 Intern. Conf. on Acoustics, Speech, and Signal Processing, 1991, p. 793 - 796.

DFKI Publikationen

DFKI Publications

DFKI Research Reports

RR-92-59
Karl Schlechta and David Makinson: On Principles and Problems of Defeasible Inheritance
13 pages

RR-92-60
Karl Schlechta: Defaults, Preorder Semantics and Circumscription
19 pages

RR-93-01
Bernhard Hollunder:
An Alternative Proof Method for Possibilistic Logic and its Application to Terminological Logics
25 pages

RR-93-02
Wolfgang Wahlster, Elisabeth André, Wolfgang Finkler, Hans-Jürgen Profitlich, Thomas Rist:
Plan-based Integration of Natural Language and Graphics Generation
50 pages

RR-93-03
Franz Baader, Berhard Hollunder, Bernhard Nebel, Hans-Jürgen Profitlich, Enrico Franconi:
An Empirical Analysis of Optimization Techniques for Terminological Representation Systems
28 pages

RR-93-04
Christoph Klauck, Johannes Schwagereit:
GGD: Graph Grammar Developer for features in CAD/CAM
13 pages

RR-93-05
Franz Baader, Klaus Schulz: Combination Tech-niques and Decision Problems for Disunification
29 pages

RR-93-06
Hans-Jürgen Bürckert, Bernhard Hollunder, Armin Laux: On Skolemization in Constrained Logics
40 pages

RR-93-07
Hans-Jürgen Bürckert, Bernhard Hollunder, Armin Laux: Concept Logics with Function Symbols
36 pages

RR-93-08
Harold Boley, Philipp Hanschke, Knut Hinkelmann, Manfred Meyer: COLAB: A Hybrid Knowledge Representation and Compilation Laboratory
64 pages

RR-93-09
Philipp Hanschke, Jörg Würtz:
Satisfiability of the Smallest Binary Program
8 pages

RR-93-10
Martin Buchheit, Francesco M. Donini, Andrea Schaerf: Decidable Reasoning in Terminological Knowledge Representation Systems
35 pages

RR-93-11
Bernhard Nebel, Hans-Juergen Buerckert:
Reasoning about Temporal Relations:
A Maximal Tractable Subclass of Allen's Interval Algebra
28 pages

RR-93-12
Pierre Sablayrolles: A Two-Level Semantics for French Expressions of Motion
51 pages

RR-93-13
Franz Baader, Karl Schlechta:
A Semantics for Open Normal Defaults via a Modified Preferential Approach
25 pages

RR-93-14
Joachim Niehren, Andreas Podelski,Ralf Treinen:
Equational and Membership Constraints for Infinite Trees
33 pages

RR-93-15
Frank Berger, Thomas Fehrle, Kristof Klöckner, Volker Schölles, Markus A. Thies, Wolfgang Wahlster: PLUS - Plan-based User Support
Final Project Report
33 pages

RR-93-16
Gert Smolka, Martin Henz, Jörg Würtz: Object-Oriented Concurrent Constraint Programming in Oz
17 pages

RR-93-17
Rolf Backofen:
Regular Path Expressions in Feature Logic
37 pages

RR-93-18
Klaus Schild: Terminological Cycles and the Propositional μ-Calculus
32 pages

RR-93-20
Franz Baader, Bernhard Hollunder:
Embedding Defaults into Terminological Knowledge Representation Formalisms
34 pages

RR-93-22
Manfred Meyer, Jörg Müller:
Weak Looking-Ahead and its Application in Computer-Aided Process Planning
17 pages

RR-93-23
Andreas Dengel, Ottmar Lutzy:
Comparative Study of Connectionist Simulators
20 pages

RR-93-24
Rainer Hoch, Andreas Dengel:
Document Highlighting —
Message Classification in Printed Business Letters
17 pages

RR-93-25
Klaus Fischer, Norbert Kuhn: A DAI Approach to Modeling the Transportation Domain
93 pages

RR-93-26
Jörg P. Müller, Markus Pischel: The Agent Architecture InteRRaP: Concept and Application
99 pages

RR-93-27
Hans-Ulrich Krieger:
Derivation Without Lexical Rules
33 pages

RR-93-28
Hans-Ulrich Krieger, John Nerbonne, Hannes Pirker: Feature-Based Allomorphy
8 pages

RR-93-29
Armin Laux: Representing Belief in Multi-Agent Worlds viaTerminological Logics
35 pages

RR-93-30
Stephen P. Spackman, Elizabeth A. Hinkelman:
Corporate Agents
14 pages

RR-93-31
Elizabeth A. Hinkelman, Stephen P. Spackman:
Abductive Speech Act Recognition, Corporate Agents and the COSMA System
34 pages

RR-93-32
David R. Traum, Elizabeth A. Hinkelman:
Conversation Acts in Task-Oriented Spoken Dialogue
28 pages

RR-93-33
Bernhard Nebel, Jana Koehler:
Plan Reuse versus Plan Generation: A Theoretical and Empirical Analysis
33 pages

RR-93-34
Wolfgang Wahlster:
Verbmobil Translation of Face-To-Face Dialogs
10 pages

RR-93-35
Harold Boley, François Bry, Ulrich Geske (Eds.):
Neuere Entwicklungen der deklarativen KI-Programmierung — *Proceedings*
150 Seiten
Note: This document is available only for a nominal charge of 25 DM (or 15 US-$).

RR-93-36
Michael M. Richter, Bernd Bachmann, Ansgar Bernardi, Christoph Klauck, Ralf Legleitner, Gabriele Schmidt: Von IDA bis IMCOD: Expertensysteme im CIM-Umfeld
13 Seiten

RR-93-38
Stephan Baumann: Document Recognition of Printed Scores and Transformation into MIDI
24 pages

RR-93-40
Francesco M. Donini, Maurizio Lenzerini, Daniele Nardi, Werner Nutt, Andrea Schaerf:
Queries, Rules and Definitions as Epistemic Statements in Concept Languages
23 pages

RR-93-41
Winfried H. Graf: LAYLAB: A Constraint-Based Layout Manager for Multimedia Presentations
9 pages

RR-93-42
Hubert Comon, Ralf Treinen:
The First-Order Theory of Lexicographic Path Orderings is Undecidable
9 pages

RR-93-43
M. Bauer, G. Paul: Logic-based Plan Recognition for Intelligent Help Systems
15 pages

DFKI Documents

D-93-01
Philipp Hanschke, Thom Frühwirth: Terminological Reasoning with Constraint Handling Rules
12 pages

D-93-02
Gabriele Schmidt, Frank Peters,
Gernod Laufkötter: User Manual of COKAM+
23 pages

D-93-03
Stephan Busemann, Karin Harbusch(Eds.):
DFKI Workshop on Natural Language Systems:
Reusability and Modularity - Proceedings
74 pages

D-93-04
DFKI Wissenschaftlich-Technischer Jahresbericht
1992
194 Seiten

D-93-05
Elisabeth André, Winfried Graf, Jochen Heinsohn,
Bernhard Nebel, Hans-Jürgen Profitlich, Thomas
Rist, Wolfgang Wahlster:
PPP: Personalized Plan-Based Presenter
70 pages

D-93-06
Jürgen Müller (Hrsg.):
Beiträge zum Gründungsworkshop der Fachgruppe
Verteilte Künstliche Intelligenz, Saarbrücken, 29. -
30. April 1993
235 Seiten
Note: This document is available only for a
nominal charge of 25 DM (or 15 US-$).

D-93-07
Klaus-Peter Gores, Rainer Bleisinger:
Ein erwartungsgesteuerter Koordinator zur
partiellen Textanalyse
53 Seiten

D-93-08
Thomas Kieninger, Rainer Hoch:
Ein Generator mit Anfragesystem für strukturierte
Wörterbücher zur Unterstützung von
Texterkennung und Textanalyse
125 Seiten

D-93-09
Hans-Ulrich Krieger, Ulrich Schäfer:
TDL ExtraLight User's Guide
35 pages

D-93-10
Elizabeth Hinkelman, Markus Vonerden,Christoph
Jung: Natural Language Software Registry
(Second Edition)
174 pages

D-93-11
Knut Hinkelmann, Armin Laux (Eds.):
DFKI Workshop on Knowledge Representation
Techniques — Proceedings
88 pages

D-93-12
Harold Boley, Klaus Elsbernd,
Michael Herfert, Michael Sintek, Werner Stein:
RELFUN Guide: Programming with Relations and
Functions Made Easy
86 pages

D-93-14
Manfred Meyer (Ed.): Constraint Processing –
Proceedings of the International Workshop at
CSAM'93, July 20-21, 1993
264 pages
Note: This document is available only for a
nominal charge of 25 DM (or 15 US-$).

D-93-15
Robert Laux: Untersuchung maschineller
Lernverfahren und heuristischer Methoden im
Hinblick auf deren Kombination zur Unterstützung
eines Chart-Parsers
86 Seiten

D-93-16
Bernd Bachmann, Ansgar Bernardi, Christoph
Klauck, Gabriele Schmidt: Design & KI
74 Seiten

D-93-20
Bernhard Herbig:
Eine homogene Implementierungsebene für einen
hybriden Wissensrepräsentationsformalismus
97 Seiten

D-93-21
Dennis Drollinger:
Intelligentes Backtracking in Inferenzsystemen am
Beispiel Terminologischer Logiken
53 Seiten

D-93-22
Andreas Abecker: Implementierung graphischer
Benutzungsoberflächen mit Tcl/Tk und Common
Lisp
44 Seiten

D-93-24
Brigitte Krenn, Martin Volk:
DiTo-Datenbank: Datendokumentation zu
Funktionsverbgefügen und Relativsätzen
66 Seiten

D-93-25
Hans-Jürgen Bürckert, Werner Nutt (Eds.):
Modeling Epistemic Propositions
118 pages
Note: This document is available only for a
nominal charge of 25 DM (or 15 US-$).

D-93-26
Frank Peters:
Unterstützung des Experten bei der Formalisierung
von Textwissen
INFOCOM - Eine interaktive Formalisierungskomponente
58 Seiten

D-94-01
Josua Boon (Ed.):
DFKI-Publications: The First Four Years
1990 - 1993
75 pages