

Endspurt für Verbomobil II – Rechner werden zum Dolmetscher

Der Computer soll Telefongespräche simultan übersetzen

Von HANS RIBBECK

Erste Spracherkennungsprogramme sind auf dem Markt. Ihre Qualität ist recht ordentlich – obwohl der Anwender noch viel Zeit investieren muß, um den Rechner an seine Stimme zu gewöhnen. In den Forschungslabors wird daran gearbeitet, auch diese Klippe zu überspringen.

HANDELSBLATT, Dienstag, 9.3.99 AACHEN. Vier Großunternehmen haben im Verbund mit Forschungspartnern 1993 ein ehrgeiziges Projekt aufgelegt: Sie wollen binnen acht Jahren eine Software für die natürliche Spracherkennung als Übersetzungshilfe beim Dialog mit ausländischen Partnern entwickeln. Vor allem Telefongespräche zwischen Deutschen und Engländern bzw. Japanern sollen in beide Sprachrichtungen gedolmetscht werden – zunächst in ausgesuchten Themenbereichen. Bei dieser Aufgabe sind die Sprachtechnologien voll im Plan. Die Entwickler wollen das Ergebnis des interdisziplinären Verbundprojekts schon im nächsten Jahr auf der Expo 2000 der Öffentlichkeit vorstellen.

Kurz vor Jahresschluß präsentierte die Philips GmbH Forschungslaboratorien – einer der insgesamt vier Industriepartner – in Aachen den Forschungsprototyp Verbomobil II als funktionierendes Dolmetschsystem für Telefongespräche. Kurz zuvor war internationalen Fachgutachtern demonstriert worden, daß das System automatisch erkennt, welche Sprache der Anrufer am Telefon spricht – ob nun Deutsch, Englisch oder Japanisch gesprochen wird. Der eigentlichen Spracherkennung haben die Forscher eine Sprachidentifizierung vorgeschaltet.

Was steckt hinter diesem fortgeschrittenen Entwicklungsvorhaben zur multilingualen Verarbeitung von Spontansprache – genannt Verbomobil? Dieses Computersystem erkennt Spontansprache sprecherunabhängig, analysiert das Gesprochene und übersetzt es in die fremde Zielsprache. Danach wird ein entsprechender Satz erzeugt und von einer synthetischen Stimme des Rechners wiedergegeben.

Es handelt sich also um eine Übersetzungshilfe beispielsweise für eine

Gesprächssituation mit ausländischen Partnern – zunächst noch für ausgewählte Themenbereiche wie die Reiseplanung, Hotelreservierung oder Terminverhandlung.

Das Projektmanagement des Verbundvorhabens liegt bei der Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH in Saarbrücken, Leiter ist Prof. Wolfgang Wahlster. In der ersten Entwicklungsphase konnte Verbomobil bereits deutsch gesprochene Eingaben aus der Domäne Terminverhandlung erkennen, analysieren und auf Basis von zunächst 1 292 Wörtern ins Englische übertragen und äußern.

Das Projekt wurde auf der CeBIT 1995 vorgestellt. Zwei Jahre später war der wiederum zur CeBIT vorgestellte Forschungsprototyp schon in der Lage, bei erweitertem Umfang von nun rund 2 500 Wörtern zusätzlich auch japanische Terminvereinbarungen ins Englische zu übersetzen.

Die 1997 begonnene zweite Phase hat für den Beginn des neuen Jahrtausends größere Schritte im Visier: Verbomobil II soll dann zusätzlich in zwei weiteren Domänen arbeiten – in der Reiseplanung und in der PC-Fernwartung. Später soll das System auf neue Gesprächssituationen einstellbar sein. Für die Reiseplanung wurde beispielsweise ein Szenario rund um die Expo als Testbereich mit bislang 7 000 Wörtern gewählt. Dafür mußte der bisherige Wortschatz allerdings verdreifacht werden.

„Wichtig ist, daß das Verbomobil für jedes dieser Wörter im Gegensatz zu reinen Diktiersystemen nicht nur die Aussprache und Schreibweise, sondern vor allem auch die Wortbedeutung im Kontext und die entsprechenden Übersetzungsvarianten kennt“, so Projektleiter Wolfgang Wahlster. Bei der PC-Fernwar-

tung, die als Anwendung für internationale Hotlines genutzt werden könnte, kommt die spezifische Fachterminologie hinzu.

Geradezu verblüffend ist, was Verbomobil heute schon leistet. Das System ermittelt bereits nach etwa sechs gehörten Worten eines Sprechers relativ treffsicher – zu 87 Prozent korrekt –, welche Sprache gesprochen wird. Diese Fähigkeit zur raschen Sprachenbestimmung ist nicht nur für das Dolmetschen internationaler Ferngespräche, sondern auch für die weitere Automatisierung multilingualer Call-Center entscheidend, weil der Anrufer nur so softwaregesteuert zu einem Mitarbeiter mit der erforderlichen Sprachkompetenz verbunden werden kann.

Das neuartige Übersetzungssystem berücksichtigt nun nach Übersetzungen von Wörtern im Satzzusammenhang auch den Dialogzusammenhang über Satzgrenzen hinweg. In der Aussage „Wir können ins Theater gehen. Es gibt eine gute Vorstellung.“ wird „Vorstellung“ durch „performance“ übersetzt. Im Dialog „Wo könnten wir uns treffen? Ich habe keine Vorstellung“ wird das Wort „Vorstellung“ treffend mit „idea“ übersetzt.

Heutige Systeme haben noch eine entscheidende Schwäche im praktischen Einsatz: Sie können nicht erkennen, ob sich ein Sprecher während einer Äußerung korrigiert hat. Selbst bei den besten Diktiersystemen erscheinen alle gesprochenen Wörter im Diktattext – unabhängig davon, ob eine Formulierung teilweise zurückgenommen wurde. Verbomobil ist das erste Dolmetschsystem, welches Korrekturen eines Sprechers erkennt und nicht mitübersetzt. So wird „Ich habe einen Termin am Montag, äh – am Dienstag“ richtig mit „I have an appointment on Tuesday“ wiedergegeben.

Zur Zeit wird intensiv daran gearbeitet, das System gegen Störungen, Fehleingaben und Ausfälle so abzusichern, daß der Dialog mit dem Benutzer nicht abreißt. Gerade bei Telefonaten ist mit Störungen des Dialogs zu rechnen. Deshalb wurde ei-



Noch ist es eine Vision, daß Computer Gespräche via Internet mit ausländischen Partner simultan übersetzen – aber die Forscher arbeiten daran.

Foto: Philips Speech Processing

ne Komponente integriert, die auch Geräusche wie Husten, Gähnen, Lachen und Räuspern erkennt und von der Spracheingabe trennt.

Völlig neu ist eine robuste Verarbeitung auf der Inhaltsebene. Werden kurze Wörter vom Sprecher schlecht artikuliert und entfallen somit bei der Worterkennung, als würde der Hörer nur Satzketten seines Dialogpartners mitbekommen, dann versucht Verbmobil, offensichtlich fehlende Wörter inhaltsorientiert so zu ergänzen, daß sich insgesamt eine sinnvolle Äußerung ergibt.

Eine weitere Entwicklungskomponente ist in der Lage, bestimmte Emotionen anhand der Stimme des

Sprechers zu erkennen. Ist der Anrufer ungeduldig oder ärgert er sich: „Was soll denn das jetzt“, schaltet das System automatisch über die Verbindung zu einem Call-Center zu einem für den Umgang mit verärgerten Kunden entsprechend trainierten Spezialisten weiter. So könne schlimmeres abgewendet werden, hoffen die Forscher. – Bei einem solchen Kunden würde die weitere Kommunikation mit dem automatischen Sprachdialogsystem eher zu weiterem Ärger führen.

Insbesondere bei der automatischen Erzeugung schriftlicher Verlaufs- und Ergebnisprotokolle von Telefongesprächen glaubt das DF-

KI-Forscherteam bald die Marktreife erreichen zu können. Im Prototyp gelang es bereits, solche Protokolle auf Deutsch oder Englisch zu versenden. Eine faszinierende Vorstellung: Man drückt nach einem Ferngespräch mit einem Geschäftspartner eine Protokoll-Taste und dem Gesprächspartner wird eine Mitschrift per Fax oder E-Mail zugestellt. Das Protokoll müßte nur noch unterschrieben werden und die Geschäftsvereinbarung wäre bereits via Telefon besiegelt.

Zum Abschluß der zweiten Phase wird das System auf einem zentralen Server verfügbar und nicht mehr von der Spracheingabe über ein Nah-

besprechungsmikrofon abhängen und selbst Freisprechen ermöglichen. Für die Spracherkennung ergibt sich damit die Herausforderung für eine Verarbeitung von Spontansprache in Telefon- oder sogar Funkqualität unter der Option des Mikrofonwechsels. Durch den Sprachserver wird Verbmobil sogar über Handy nutzbar.

Im Verbundvorhaben Verbmobil kooperieren mit der Daimler-Chrysler AG, der Philips GmbH und der Siemens AG große Unternehmen der Informationstechnologie. Zählt man Universitäten und Forschungszentren mit, sind insgesamt 23 Projektpartner mit von der Partie. Projektträger ist die Deutsche Forschungsanstalt für Luft- und Raumfahrt (DLR). Das BMBF fördert das Projekt bis zum Jahr 2000 mit 115 Mill. DM, die Industriepartner beteiligen sich mit 51 Mill. DM. Die Teilergebnisse stehen allen Unternehmen, insbesondere also auch dem Mittelstand, zur Verfügung; Lizenzen werden über die Konsortialpartner abgegeben.

Der Forschungsprototyp Verbmobil II wird aus 59 Sprachverarbeitungsmodulen bestehen, die über 131 Datenpools teilweise als parallele Prozesse miteinander kommunizieren. Zusammengeführt werden sie in einer streng objektorientiert realisierten Software-Architektur zu einem komplexen Übersetzungssystem. Dabei wird ein Softwarepaket entstehen, das an die Komplexitätsgrenzen dessen stößt, was derzeit in der Sprachtechnologie machbar ist.

DFKI-Gesamtprojektleiter Wahlster ist sich sicher: „Es gibt weltweit kein Dolmetschsystem, das mehr Funktionsmerkmale aufweist als Verbmobil II“. Aber noch ist zusätzliche Detailarbeit nötig: Die Module müssen noch besser aufeinander abgestimmt werden und die Laufzeit muß verkürzt werden, Robustheit und sprachliche Abdeckung sind weitere Herausforderungen. Besonders, wenn Verbmobil nicht nur in Dialogsituationen, sondern auch in Telekooperationsanwendungen mit mehreren Gesprächspartnern Übersetzungshilfe leisten soll.