

FORUM

Alexander Waibel

Die Technik hilft, Brücken zu schlagen

Immer mehr wird die Technik die Kommunikation zwischen Menschen und das Abliefern bzw. Abfragen beliebiger gespeicherter Informationen unterstützen. Dabei müssen verschiedene Barrieren – nicht zuletzt die Sprachbarrieren – überwunden werden. Die Arbeit bei der Schaffung automatischer Sprachübersetzungssysteme und sinnfälliger Mensch-Maschine-Schnittstellen hat in letzter Zeit große Fortschritte gemacht. Viel ist bereits erreicht worden, doch viel Arbeit bleibt auch noch zu tun, um den Menschen die Chance zu geben, weltweit mit Hilfe der Technik einander näher zu kommen.

Die neue Qualität der Informationstechnik

Der Computer steht in einem rasanten Wandel – vom Rechner als Großanlage, der in abgeschotteten Rechenzentren vor sich hinratterte, hin zum Haushaltsgerät und Gebrauchsgegenstand, der für ein paar hundert Mark ein Vielfaches an Rechenleistung liefert. Dazwischen liegen nur zwei Jahrzehnte. Und doch ist dies erst der Anfang. Ein Blick in die Zukunft ergibt, daß uns noch vor der Jahrhundertwende der PC im Gigaflop-, Gigahertz- und Gigabyte-Bereich zu niedrigen und erschwinglichen Kosten beschert wird. Wenn nun solche enorme Rechenleistung einer solchen Bevölkerungsbreite verfügbar gemacht werden soll, so tritt das eigentliche „Rechnen“ – ja der Anschein des Technischen – immer mehr in den Hintergrund. Der einfache Zugriff, das spielerische, menschlich nahe Benutzen ohne Lehrbücher, Manuale und Knöpfe, wird zum Ziel der Entwicklung, und es werden 90 % der Rechenleistung hierfür in Beschlag genommen.

Die Grenzen und Definitionen traditioneller Wirtschaftsbereiche, wie Kommunikation, Unterhaltung und Computing, fangen damit auch an, sich zu verschieben und zu verschwimmen. Wenn Microsoft-Gründer Bill Gates heute von einem weltumspannenden Satellitenfunknetz träumt, und das kommende Jahrzehnt deren Realisierung bringt, beschleunigt dies das Verschwinden nationaler Grenzen. Kommunikation zwischen Menschen, Herbeirufen, Abliefern und Abfragen beliebiger Informationen, in beliebiger Form,

zu beliebiger Zeit, an beliebigem Ort, durch einen beliebigen Anbieter wird alltäglich und preiswert sein, und zunehmend wird die Information und deren Verarbeitung (die Software) im Vordergrund stehen. So werden Dienstleistungen und geistige Kreativität weltweit verfügbar – ein positiver Trend, wenn man an kulturellen Austausch und bessere Völkerverständigung denkt und auf eine größere Betonung geistiger und kultureller Werte in unserer technokratischen und formalisierten Welt hofft. Enorme Freiheit und Beweglichkeit wird sich entfalten.

Defizite an der Mensch-Maschine-Schnittstelle

Erschreckend sind in dieser Vision bisher nur die Ein- und Ausfahrten solch weltweiter Informations-Superhighways. Was kann der einzelne anbieten, wie kann er mithalten, ist der Zugriff menschengerecht und menschlich? Und hier ist noch manch ein Traum zu träumen. Barrieren zwischen Mensch und Maschinen entstehen:

- Zum einen wegen der Flut von Informationen: Wer kann noch 500 Fernsehkanäle überblicken und einordnen?
- Zum zweiten entstehen Barrieren aus mangelndem Wissen, die Technik und ihre vielfältigen Möglichkeiten bedienen und kontrollieren zu können. Schon heute können die wenigsten ihren Videorecorder programmieren, geschweige denn komplexe Informationen auf einer weltweiten „Infobahn“ interaktiv abrufen.

Auch existieren natürliche Grenzen



Prof. Dr. Alexander Waibel (38) ist Informatiker an der Universität Karlsruhe und beschäftigt sich vor allem mit der Mensch-Maschine-Kommunikation und mit Neuronalen Netzen. Im Rahmen einer Kooperation zwischen der Carnegie Mellon University (USA) und der Universität Karlsruhe ist unter seiner Leitung das Sprachübersetzungssystem „Janus“ entwickelt worden, das gesprochene Sprache simultan vom Deutschen oder Englischen in mehrere Zielsprachen übersetzt. Für seine Arbeiten an automatischen Übersetzungssystemen erhielt er den diesjährigen, mit 35 000 DM dotierten Forschungspreis „Technische Kommunikation“ der Alcatel-SEL-Stiftung. Der Beitrag ist eine leicht gekürzte Fassung seiner Rede anlässlich der Preisverleihung.

zwischen Menschen wie physische Distanz, begrenzte Lernfähigkeit, Behinderung, kulturelle Verschiedenheit und nicht zuletzt Sprachbarrieren. Das Überwinden dieser Barrieren bedeutet nicht nur, menschlichere Technik zu schaffen, sondern auch in einer Marktwirtschaft strategische Vorteile und damit Arbeitsplätze in einer Informationsgesellschaft zu sichern.

Sprachbarrieren müssen überwunden werden

Aber wie können diese Barrieren überwunden werden? In der Tat sind viele der zuvor genannten Schwierigkeiten Sprachbarrieren, Barrieren zunächst zwischen Menschen, die verschiedene menschliche Sprachen sprechen, Barrieren aber auch zwischen einem Mensch, der nur eine vague, unsichere, gestotterte Frage stammelt, und einer Maschine, die (trotz Fuzzy-Logik) allgemein nur einen präzise formulierten Code und starre, computergerecht eingetippte Anfragen toleriert.

Das erste Problem ist vielleicht das schwierigste, gleichzeitig aber auch eines der aufregendsten: Wie können ein Deutscher und ein Kollege in Japan an einem virtuellen Videokonferenztisch zusammensitzen, miteinander sprechen, ohne eine gemeinsame Sprache zu beherrschen? Der Traum der weltweiten Verständigung soll in unserer Generation noch realisierbar sein. Ein übersetzendes Telefon, eine multinationale, multilinguale, multikulturelle Videokonferenz, automatische Untertitel bei ausländischen Nachrichtensendungen, der Videoein-

kaufsbummel und das Handeln mit dem Anbieter in Fernost usw. – ein weltweiter Horizont eröffnet sich. Wie die Gutenberg-Pressen das geschriebene Wort popularisierte und demokratisierte, soll nun das ausländische gesprochene Wort einem jeden zugänglich werden. Warum sehen wir uns aber heute imstande, einen solchen Traum zu träumen, und wie realistisch ist er wirklich?

Von der Mehrdeutigkeit gesprochener Sprache

Zunächst ist wichtig zu verstehen, wie mehrdeutig die Welt der Signale um uns herum ist und mit welcher erstaunlichen Leichtigkeit der Mensch sich aus der Flut der möglichen Interpretationen des Gesprochenen, Geschriebenen und Gesehenen etwas Passendes heraus sucht. Im gesprochenen Sprachsignal liegen dabei zunächst viele akustische Störungen vor. In einem Saal voller Menschen kann man ohne weiteres dem Redner folgen, auch wenn der Nachbar hustet, der Vordermann etwas flüstert und der Raum mit Echo und Hall Interferenzen produziert. Man versteht den Redner auch dann, wenn er ins Stocken oder Stottern kommt. Aus einem „sollen wir einmal“ nach Duden wird tatsächlich ein „sollmama“, und fast in jeder akustischen Sequenz verbirgt sich eine alternative akustisch richtige (wenn auch im Kontext nicht intendierte) Interpretation. Dies ist übrigens in jeder Sprache so. Im Englischen ist z. B. „Give me a New Display“ und „Give me a Nudist Play“ akustisch nicht differenzierbar.

Mehrdeutigkeit wird vom Menschen mit großer Leichtigkeit aufgelöst, ja das Aufdeuten der weniger wahrscheinlichen Alternativen ist meist verblüffend und oft die Würze und Pointe eines Witzes.

Das zweite große Problem bei der Übersetzung gesprochener Sprache ist die Verarbeitung und Übersetzung der Sprache (wenn sie erst einmal akustisch erkannt ist) von einer Landessprache in die andere (sagen wir z. B. vom Deutschen ins Japanische). Auch hier bildet Mehrdeutigkeit den Kern der Problematik. Aus informationstheoretischer Sicht kann sie auch hier wie eine Geräuschkulisse, nur auf semantischer Ebene, gedeutet und berechnet werden. Vielleicht kennen manche von ihnen die berühmte Anekdote in der maschinellen Übersetzung, in der der biblische Satz: „The Spirit is willing but the flesh is weak“ als „The vodka is good, but the meat is rotten“ übersetzt wurde.

In der Übersetzung gesprochener Sprache kommen nun auch noch die Fehler in der akustischen Erkennung hinzu, und – wenn wir nun echte, menschliche Dialoge in realen Situationen zulassen wollen – die Sprechfehler spontaner Sprache und wiederholt auftretende Geräusche des Alltags (Schmatzen, Räuspern, Lachen, Telefonklingeln, Türschlagen usw.).

Heutiger Stand und nächste Ziele der automatischen Sprachübersetzung

Unser erstes System „Janus“ hat mit einem begrenzten Vokabular in einer begrenzten Gesprächsdomäne

und nur unter Beachtung syntaktisch korrekter (vorgelesener) Sätze seine „Jungfernfahrt“ vor einem Jahr erfolgreich bestanden. Man konnte sich damit nur zu einer Konferenz in Japan anmelden und mußte sich ohne Husten und ohne Schmatzen an nur 500 zulässige Wörter halten. Immerhin konnte es aber zum ersten Mal kontinuierliche Sätze sprecherunabhängig erkennen, analysieren und in eine Zielsprache hinein fast immer fehlerfrei übersetzen. In einer Videokonferenzschaltung wurde es zusammen mit Systemen bei Siemens und ATR in Japan im Januar 1993 der Öffentlichkeit vorgestellt.

Ein Jahr ist seitdem ins Land gegangen, und unser Team arbeitet nun an der schwierigsten Problematik – der Spontansprachlichkeit. Zwar wird auch hier noch eine Einschränkung der Domäne zugrunde gelegt, aber es wird einem Gesprächspartner nicht mehr das Korsett eines eingeschränkten Vokabulars oder einer eingeschränkten Sprechweise vorgeschrieben. Er soll nun Husten, Schmatzen und „Äh“ sagen dürfen, er soll mitten im Satz abbrechen können, und sein Satz soll – nicht wie bei geübten Fernsehansagern – auch einmal grammatikalisch falsch sein dürfen und trotzdem verstanden werden. Diese Problematik erfordert eine stärkere Betonung der Semantik, d. h. des Verstehens des Inhalts, gegenüber der reinen Analyse der Struktur (Syntax) einer Äußerung. Auch das wird in diesen Tagen tatsächlich Realität. Unter Zu-

hilfenahme einer semantischen Darstellung der relevanten Bedeutungen in einer Diskursdomäne sehen wir uns nun imstande, die wichtigen Inhalte aus einer Äußerung zu extrahieren und diese dann in beliebigen Sprachen wiederzugeben. Diese Vorgehensweise hat zwei Vorteile: Erstens wird die Erweiterung auf mehrere Sprachen erleichtert und kostet nur noch geringen Zusatzaufwand (wir arbeiten bereits an Deutsch, Japanisch, Englisch, Spanisch und Koreanisch). Zweitens wird eine sinngemäße, ja kulturell angepaßte Übersetzung ermöglicht. Das System kann den Sinn eines Ausdrucks in der einen Sprache durchaus mit einer äquivalenten Redewendung in der anderen wiedergeben. So kann z. B. ein „geht nicht“ oder „unmöglich“ im Deutschen, je nach Kontext, durchaus mit dem im Japanischen adäquateren „Muzukashii“ oder „schwierig“ wiedergegeben werden. Auf der Seite der Zielsprache muß dann letztlich auch noch ein akustisch hörbarer Satz ausgegeben werden. Es gibt hierfür bereits eine Reihe kommerzieller Systeme, die aber noch in unserer Zukunftsvision an die Stimme des Eingabesprechers angepaßt werden müssen.

Effektive Kommunikation zwischen Mensch und Maschine

Soweit also das Brückenschlagen auf der „Infobahn“. Wie steht es aber mit den gefährlichen Ein- und Ausfahrten? Glücklicherweise ist auch

hier wieder dieselbe Technologie, wenn auch mit anderen Vorzeichen, im Spiel. Zur effektiven Kommunikation zwischen Mensch und Maschine sollten nicht komplizierte Codewörter und Programmsequenzen vonnöten sein, sondern auch die gestammelte Anfrage des Ungeübten soll verstanden werden. Die Maschine soll nicht transkribieren, sondern verstehen und assistieren. Genau in diese Richtung entwickelt sich auch unsere Arbeit. Wieder sind semantische Rahmen eine Ausdrucksform des Inhalts, und Wortsequenzen und Fragmente werden auf diese Rahmen zum Verständnis abgebildet.

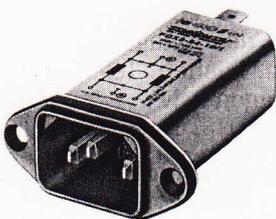
Menschliche Kommunikation ist nun zu guter Letzt auch nicht immer nur gesprochene Sprache. Man nutzt auch Körpersprache, Augenkontakt, Gesichtsausdruck, Ansprache, Zeigegestik, Lippenbewegung und Handschrift als alternative Äußerungsmodalitäten. Im Labor haben wir begonnen, diese „multimodalen“ Kommunikationssignale systematisch anzugehen und auch auf ihren Inhalt hin zu analysieren und semantisch abzubilden.

Wenn man Systeme realisieren möchte, die eine natürliche Interaktion zwischen Mensch und Maschine bieten und zwischen Mensch und Mensch vermitteln und interpretieren können, muß auch hier eine bessere Anpassung der Maschine an den Menschen das Ziel sein.

Nun stehen wir mit unseren Arbeiten noch ganz am Anfang einer Entwicklung, bei der noch manch „harte wissenschaftliche Nuß“ geknackt werden muß. Gesellschaftlich aber ist jetzt schon klar, daß die Umbrüche in Technik und Kommunikation Welt und Menschheit einander näher bringen werden, daß dadurch die Zukunft nur eine gemeinsame Zukunft des weltweiten Zusammenwirkens sein kann. Dies wird enorme Chancen für Produkte, Dienstleistungsangebote und Märkte bedeuten, aber auch enorme Anpassungsbereitschaft, schnelleres Handeln und Reagieren, Flexibilität und Verständnis erfordern. Ein Umdenken inflexibler Strukturen, ein neues Selbstverständnis von „Management“ und das Heranbilden von dynamischen und internationalen gewandten Fachkräften ist somit die Herausforderung des kommenden Jahrzehnts.

TIMONTA

Steckerfilter für höchste Sicherheit



TIMONTA konzipierte eine neue Filterfamilie mit IEC-Kaltgerätestecker in den Stromstärken 1 A, 2 A, 4 A, 6 A und 10 A.

Gute Dämpfungseigenschaften, verstärkte Isolation, X1 / Y-Kondensatoren, selbstverlöschende Kunststoffe nach UL und die Berücksichtigung der Anforderungen nach EN 60950 und IEC 664 garantieren hohe Sicherheit.

Für normale Anforderungen ist eine Low-cost-Version im verkürzten Gehäuse lieferbar.



Bewilligung erteilt

TIMONTA GmbH + Co.

79108 Freiburg i.Br. • Wöhlerstraße 1-3
Tel. (07 61) 50 41 50 • Fax (07 61) 50 21 87