

Computer zeigen Sprachgefühl

Verständnis auch bei doppeldeutigen Formulierungen / Übersetzung ohne menschliche Hilfe

Eine Maschine, die Sprache „versteht“ und sinnvoll antwortet, versuchen Forscher schon seit 30 Jahren zu entwickeln. Wie schwierig diese Aufgabe ist, zeigt sich daran, daß erst jetzt die ersten Spracherkennungssysteme auf den Markt kommen. Sie erfüllen einfache Aufgaben wie das Weiterleiten von Telefonanrufen, sie regeln Bankgeschäfte oder nehmen Bestellungen auf. Eine automatische Fahrplanauskunft der Firma Philips in Aachen wird zur Zeit erprobt. Die ersten Werkzeugmaschinen und Personal-Computer, die sich durch Sprache steuern lassen, sind bereits auf dem Markt. Die von Daimler-Benz entwickelten sprachgesteuerten Autos stehen kurz vor der Markteinführung. Philips bietet ein Spracherkennungssystem an, das diktierbare radiologische Befunde aufschreibt, während IBM ein ähnliches System für Geschäftsbriefe entwickelt hat.

Die zahlreichen Möglichkeiten, die Spracherkennung zu nutzen, sind allerdings noch lange nicht ausgeschöpft. Deutsche Forscher streben auf diesem Gebiet eine Spitzenposition an. Vor drei Jahren haben sich sieben Industriepartner und 22 Universitäten und Forschungszentren zusammengeschlossen, um ein anspruchsvolles Ziel zu verwirklichen: einen Übersetzungcomputer, der den Verlauf des Gesprächs verfolgt und bei Unklarheiten sogar zurückfragt. Das umfangreiche Projekt mit dem Namen „Verbmobil“ wird in der ersten Phase, die bis Ende 1996 dauert, vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie mit insgesamt 64,9 Millionen Mark unterstützt. Die Industriepartner finanzieren das Projekt mit weiteren 31 Millionen Mark.

Der Markterfolg von Spracherkennungssystemen hängt entscheidend davon ab, inwieweit sie spontan gesprochene Sätze unabhängig vom Sprecher „verstehen“ können. Damit sind zahlreiche Schwierigkeiten verbunden: Es gibt nicht nur alte und junge Stimmen, sondern auch unterschiedliche Dialekte und Sprachmelodien. Selbst wenn derselbe Mensch zweimal denselben Satz spricht, registriert der Rechner leicht voneinander abweichende akustische Schwingungen. Darüber hinaus sprechen die wenigsten Menschen schriftreif. Sie un-

terbrechen oder räuspern sich, niesen, bauen Füllwörter ein, wiederholen ihre Aussage in anderen Worten oder brechen Sätze ab mit dem Verweis: „Du weißt schon, was ich meine...“ Eine solche Formulierung, die bereits bei menschlichen Gesprächspartnern mit unterschiedlichen Erfahrungen und Denkweisen für Mißverständnisse sorgt, überfordert den Computer erst recht. Außerdem fehlt ihm die Möglichkeit, zusätzliche Informationen über die Mimik, Gestik und emotionelle Klangfärbung der Stimme zu gewinnen.

Am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Saarbrücken, das die Federführung des „Verbmobil“-Projekts übernommen hat, erforschen Wissenschaftler die Tücken der Sprache genauer. Ihr Ziel ist ein Sprachverarbeitungssystem, das deutsche und englische Gesprächspartner bei der Absprache von Terminen unterstützt. Um dies erreichen zu können, zeichneten sie bis jetzt mehr als 2000 Dialoge zwischen Geschäftspartnern auf Tonband auf. Sie stellten beispielsweise fest, daß die Satzmelodie eine große Hilfe bietet, den Sinn doppeldeutiger Sätze zu erkennen. Die Aussage „Ja-zur-Not-geht-es-auch-Samstag“ kann interpretiert werden als „Ja, zur Not geht es auch Samstag“ oder „Ja, zur Not! Geht es auch Samstag?“ Damit auch „Verbmobil“ diese feinen Unterscheidungen treffen kann, bauten die Wissenschaftler unter Leitung von Wolfgang Wahlster ein sogenanntes Prosodiemodul ein. Es registriert die Pausen zwischen den Wörtern, den Tonhöhenverlauf sowie Dauer und Energie des Signals, womit sich Satzanfänge und -enden erkennen lassen.

Darüber hinaus braucht das System Regeln über den Satzbau und die Bedeutung von Wörtern in unterschiedlichen Zusammenhängen. In der Aussage „Es geht bei mir“ beispielsweise kann „bei mir“ einen Ort oder einen Termin bedeuten. Unvollständige Sätze kann „Verbmobil“ aufgrund der gelernten Regeln selbst ergänzen. Eine weitere Hilfe bietet das sogenannte Dialogmodul. Es unterteilt den Dialog in verschiedene Akte, beispielsweise das Annehmen oder Ablehnen eines Termins. Mit Hilfe von statistischen Methoden ermittelt es

dann, um welchen Dialogakt es sich handelt und welcher folgen wird. So „weiß“ es beispielsweise, daß in dem Satz „Ja, da habe ich keine Zeit“ das Wort „ja“ als Füllwort dient und nicht mit „yes“, sondern mit „well“ übersetzt werden muß.

Der größte Fortschritt besteht darin, daß „Verbmobil“ selbst die Äußerungen von unentschlossenen und verwirrten Gesprächspartnern „versteht“. Die Universität Karlsruhe, die am „Verbmobil“-Projekt beteiligt ist, hat eigens „Müllmodelle“ entwickelt, die das Gesagte von überflüssigen „Ehms“, „Ähs“, Räuspern und Schmatzen trennen. Bei Siemens wurde darüber hinaus ein System zur robusten Analyse entwickelt, das fehlerhafte Äußerungen selbständig korrigiert. Eine weitere Hilfe ist die sogenannte flache Verarbeitung, bei der in unzusammenhängenden Äußerungen nach Schlüsselwörtern gesucht wird. Mit der Wahrscheinlichkeitsrechnung ermittelt das System dann, was der Sprecher hätte sagen wollen. Künftige Entwicklungen sollen es ermöglichen, daß der Computer in einem solchen Fall zurückfragt, ob seine Interpretation richtig ist.

Bisher beherrscht der Verbmobil-Demonstrator einen Wortschatz von etwa 1300 deutschen Wörtern und deren englische Übersetzung. Ein Forschungsprototyp, der bis Ende 1996 fertiggestellt sein soll, umfaßt bereits 2300 Wörter und erkennt auch japanische Eingaben. Bei dieser Sprache kommt die Schwierigkeit hinzu, daß die Formulierungen sich beispielsweise danach richten, ob der Gesprächspartner männlich oder weiblich, sozial höher oder niedriger gestellt ist. Die zukünftigen Entwicklungen sollen auch weitere Sprachpaare in die Übersetzung einbeziehen und das Vokabular auf andere Gesprächsthemen erweitern. Auch für Übersetzungen über Telefon, bei denen die Qualität der Sprachsignale deutlich schlechter ist, soll „Verbmobil“ künftig verwendet werden. Der Aufwand erscheint aus Sicht der Marktforscher äußerst lohnend. Sie sagen der Sprachverarbeitungstechnik jährliche Wachstumsraten von weltweit 21 Prozent und europaweit sogar 32 Prozent voraus.

ANNE HARDY