
Man kann sich vorstellen, wie Microsofts PR-Leute, Entwickler und Wissenschaftler hinter den Kulissen in die Tischkante bissen. Eigentlich hatten sie ja vorführen wollen, wie Computer endlich die letzte Barriere für die Völkerverständigung niederrichten: Nachdem die Menschheit global per Telefon und Internet vernetzt ist, sollen Rechner die Grenzen zwischen den Tausenden verschiedener Sprachen aufheben. Stattdessen geriet ihre Show zu einer Demonstration der Schwierigkeiten, die diesem Menschenheitsraum in Detail entgegenstehen.


Aus der Science-Fiction kennen wir das schon lange, etwa aus Douglas Adams' Roman Per Anhalter durch die Galaxis. Da stecken sich die Leute den sogenannten Babelfisch ins Ohr, der alles Fremdländische in die eigene Sprache dolmetscht.

»Mir geht es gut, wie geht es dir?« wurde plötzlich zu: »Well how is it me friends?«

Der Protokollroboter C-3PO aus Star Wars beherrscht sechs Millionen verschiedene Sprachen. Jetzt wird der jahrzehntelang Traum Wirklichkeit, umsonst und für eine breite Masse: 300 Millionen Menschen nutzen den (in der Basisversion kostenlosen) Dienst weltweit. Wie die maschinelle Simultanübersetzung in den Alltag tritt und wie sie funktioniert, das sagt viel über ein Forschungsfeld aus, das wie kein zweites für die Hybris der Informatik steht: KI, die Künstliche Intelligenz.


So eine Face to Face-Ubersetzung ermöglicht Apps wie Google Translate oder Jibbigo, eine zum Teil in Deutschland entwickelte Software, die inzwischen von Facebook gekauft wurde. Deren Güte ist akzeptabel: Sie missverstehen vieles, produzieren unfreiwillige Konvik, aber der Sinn des Gesprächen kommt meistens rüber.

Allerdings spricht der Nutzer bei solchen Anwendungen mit einem Satz in den Computer und wartet auf die Übersetzung. Und weil er weiß, dass eine Maschine zubehört, spricht er besonders deutlich und korrekt. »In einer Konversation ist das ganz anders«, erklärt Chris Wendt, bei Microsoft Programm-Manager fürs maschinelle Übersetzen.

»Da gibt es falsche Anfänge, Stotterer, was ich noch, ähm, was ich noch, ähm, sagen wollte, ähm, kleine und größere Versprecher – all das will man nicht in der Übersetzung hören.«

Um Gespräche übersetzen zu können, braucht es eine Software drei Teilprogramme: Das erste extrahiert aus dem Tonsignal zuerst Laut und ordnet diese Buchstaben und Wörter zu. Das zweite übersetzt dieses schriftliche Ergebnis in die Fremdsprache. Und das dritte liest das Ergebnis der Übersetzung vor.


Praktisch scheiterten die Systeme lange an der schwierigsten Technik: Doch inzwischen existieren Chips, die mit 1000 Prozessoren wirklich große Netze mit vielen Zwischenschichten simulieren können (Deep Neural Networks). Microsoft gelang es, mit so einem System die Zahl fälschlicher Wörter bei spontan Gesprochenem von 30 auf 20 Prozent zu senken.


Besonders faszinierend findet er, dass die Netze offenbar sehr ähnlich lernen wie ein Mensch. Obwohl man ihnen nur den Input und die gewünschten Output präsentiert, spezialisieren sich einzelne Neuronen bald auf bestimmte Details, etwa die kleinsten Lauteinheiten (Phoneme).


Während aber die neuronalen Netze bei der Spracherkennung immer noch besser werden, stagniert der Fortschritt beim Übersetzen. Mehr Trainingsbeispiele machen die Statistik nicht mehr viel besser. Wahrscheinlich müssen auch beim Übersetzen bald die tiefen Netze ran – nur weiß wohl noch niemand so richtig, wie.

Ein weiteres Problem ist, dass diese Übersetzungsverfahren nicht mit allen Sprachen gleich gut zurechtkommen. »Wenn ich ein Demo machen sollte, würde ich mir auch Englisch/Spanisch aussuchen.« So erklärte Alex Waibel, warum Microsoft als Debut des Skype-Dolmetschers

Aber Microsoft will seine Probleme mit dem Deutschen bald beheben. Zwar legt sich Chris Wendt auf kein Datum für einen deutschen Skype-Übersetzer fest, verspricht aber: »Man kann dieses Jahr damit rechnen.« Noch läuft die Dolmetscherfunktion in einer öffentlichen Testphase. Um sie kontinuierlich zu verbessern, setzt Microsoft auf menschliche Helfer. Transkripte müssen überprüft und bewertet werden; dazu braucht man Kräfte, die beide Sprachen beherrschen. Wo die sitzen? Microsoft-Mann Wendt antwortet ausweichend: »Die finden Sie nicht in Kalifornien, in Deutschland oder in Japan.« Sondern? Es gebe Firmen, die so etwas zu einem sehr günstigen Preis machten. Man darf vermuten, dass diese polyglotten Telearbeiter in Ostasien zu finden sind.


Microsoft hat die Computerübersetzung nicht neu erfunden, traut sich aber als Erster mit seinem Babelfisch für die Internettelefonie an die schwierigsten aller Gesprächssituationen heran: die ganz normale Unterhaltung. Weil die Firma es an dem weitgehend kostenlosen Dienst Skype probiert, verringert sie die Sprachbarrieren zwischen Menschen in aller Welt. Und sie trägt dazu bei, dass die lange geschrumpfte KI wieder Öffnung bekommt. Viele Experten glauben, dass nun tatsächlich das Zeitalter der klugen Computer anbricht. »Diese Netze kommen dem Traum von intelligenten Maschinen näher«, sagt der Veteran Alex Waibel. Die Bildqualität von Skype beim Interview ist leider zu schlecht, um zu sehen, ob er dabei Tränen in den Augen hat.