

Dieter E. Zimmer
Braem-Preis-Rede
Bergneustadt 23.11.1996

Lieber Herr Ingendaay,
liebe Kolleginnen und Kollegen,

ich glaube, ich sollte in diesem Augenblick nicht nur mit vielen Worten danke sagen, schon weil ich mir dann vorkäme, als befände ich mich mitten in Christoph Marthalers „Murx“, wo sie so ergreifend singen: „Danke für alle guten Freunde. Danke, o Herr, für jedermann. Danke, wenn auch dem größten Feinde ich verzeihen kann. // Danke für diesen guten Morgen. Danke für diesen schönen Preis. Danke, daß ich mich aller Sorgen nunmehr ledig weiß.“

So also nicht. Ich sollte Ihnen vielmehr etwas übers Übersetzen sagen. Aber kann man Übersetzern etwas übers Übersetzen sagen, was sie nicht längst schon wissen? Auch auf die Gefahr hin, daß ich niemandem hier etwas Neues verrate, will ich ein paar Anmerkungen zu einem meiner Lieblingsthemen machen, der digitalisierten Sprache.

Der Ausgangspunkt soll eine simple Feststellung sein, nämlich: Wir können übersetzen, aber der Computer kann nicht übersetzen. Entgegen allem, was uns seit fünfzig Jahren immer wieder versprochen wird, kann der Computer noch nicht übersetzen und wird es in absehbarer Zeit auch nicht lernen.

Das sage ich ganz ohne die übliche erleichterte Schadenfreude: Ätsch, kann er nicht! Sie ist schon etliche Male zuschanden geworden. Es hieß auch einmal, der Computer werde nie Schach spielen oder gesprochene Sprache „verstehen“ können, und dann konnte er es doch, erst einigermaßen und schließlich besser als die meisten Menschen. Der Test wäre immer noch der, den sich Alan Turing in den fünfziger Jahren ausgedacht hat. Man stelle einen Computer in einen Raum, setze einen Menschen in einen anderen, in einem dritten nehmen die Frager Platz und stellen, zum Beispiel über eine Tastatur, ihre Testfragen. Sie wissen

nicht, wer ihnen antwortet, der verborgene Mensch oder der verborgene Computer. Und wenn der Computer ihnen gleich gute Antworten gibt wie der Mensch, sie den einen also nicht mehr von dem andern unterscheiden können, dann hat der Computer den Test bestanden. Sie sprechen also zum Beispiel ein irgendein unerwartetes Wort ins Mikrofon, etwa „Wermutstropfen“, und wenn es kurz darauf richtig ausgeschrieben auf ihrem Bildschirm erscheint und es nicht der Mensch, sondern der Computer hingeschrieben hat, dann kann man sagen, der Computer habe es „verstanden“, und wenn er auf beliebige Wörter richtig reagiert: er verstehe Sprache. Es ist klar, daß Computer heute viele solcher Tests glänzend bestehen würden. Beim Schachspielen würden sie sich am ehesten dadurch verraten, daß sie so gut spielen.

Beim Übersetzen aber würde heute kein Computer den Turing-Test bestehen. Auch ein Mensch, der kein Übersetzer ist, würde ihn spielend schlagen. Das ist für uns beruhigend, aber warum ist das eigentlich so?

Ich denke überhaupt nicht gering von den Fähigkeiten des Computers. Ich meine auch nicht, daß er uns bei unserer Arbeit nicht von großem Nutzen sein kann. Unsere Sorgen von vor zehn Jahren haben sich von selbst erledigt. Damals fürchteten wir, die Verlage würden uns zwingen, unsere Arbeit auf Disketten abzuliefern, und wir hätten die unangenehme Mehrarbeit und sie den Vorteil. Dann kam es oft genau andersherum: Wir wollten Disketten abliefern, schon um uns die bevorstehende Arbeit des Korrekturlesens besonders bei Registern, Bibliographien und so weiter zu erleichtern, aber die Texterfassungsbüros der Verlage stellten sich verzweifelt ungeschickt an beim Umgang mit unseren Disketten. Der Computer ist inzwischen wahrscheinlich für die meisten von uns ein ganz unentbehrliches Werkzeug geworden, und wir würden auf ihn auch dann nicht verzichten, wenn die Verlage uns die vorweggenommene Texterfassungsarbeit gar nicht honorierten. Wir können uns auch vorstellen, wie er sich noch nützlicher machen könnte, zum Beispiel durch die Einbindung wirklich guter großer Wörterbücher in unsere Textverarbeitungen. Der Nutzen bestünde vor allem darin, daß wir, wenn wir irgendwo unterwegs arbeiten wollen, nicht immer einen großen schweren Koffer mit

Nachschlagewerken mitschleppen müßten, also in einem erheblichen Mobilitätswachstum. Schon heute stecke ich mir die ganze *Encyclopaedia Britannica* in die Jackentasche, wenn ich mit meinem Notebook ins Grüne oder ins Blaue fahre. Sogar Elmar Tophovens großer Traum könnte dank Computer eines Tages Wirklichkeit werden: zweisprachige Wörterbücher, in denen steht, was in normalen Wörterbüchern nicht steht, nämlich was bei schwierigen Wörtern und Wendungen Übersetzerkollegen sich haben einfallen lassen. Ich selber würde manchmal sehr gerne irgendwo nachschlagen können, wie andere Übersetzer so simple englische Adjektive wie *prim* oder *bland* oder *contingent* in verschiedenen Zusammenhängen übersetzt haben.

Darum also handelt es sich nicht: um die Nützlichkeit des Computers als Arbeitswerkzeug. Das kann er, und zwar immer besser. Aber selber übersetzen: das kann er nicht. In diesem Jahr kann sich jeder zum ersten Mal selber davon überzeugen. Zwei fortgeschrittene, hochraffinierte Übersetzungsprogramme wurden für den PC zurechtgeschneidert, das eine von IBM entwickelt (*PT*), das andere von Siemens (*METAL* alias *T1*). Bis vor kurzem kostete dieses noch 60.000 Mark und ein paar Wochen Einarbeitungszeit, jetzt kann man es für unter 600 Mark haben und braucht es sich nur noch von der CD-Rom auf die Festplatte zu laden. Ich will kurz zitieren, was beide aus dem ersten Absatz von Nabokovs „Sebastian Knight“ gemacht haben:

PT: Sebastian Ritter wurde auf der dreißig geboren-erst vom Dezember, 1899 in der früheren Hauptstadt meines Landes. Eine alte russische Dame, die aus irgendeinem obskuren Grund hat, bat mich, nicht ihren Namen zu eröffnen, zeigte mich zufällig in Paris das Tagebuch, das sie in der Vergangenheit gehalten hatte. So ereignislos waren jene Jahre (anscheinend) gewesen, daß das Sich Ansammeln von täglichen Details (das immer eine arme Methode für Selbst Preservation ist) kaum eine kurze Beschreibung vom Wetter des Tags übertraf; Und es ist neugierig, in dieser Hinsicht zu beachten, daß die persönlichen Tagebücher von Herrschern, nein, wichtig sind, was Schwierigkeiten suchen heim ihre Reiche kümmern sich hauptsächlich um dasselbe Thema.

T1: Sebastian-Ritter war auf dem dreißigersten von Dezember 1899 geboren, in der früheren Hauptstadt meines Lands. Eine alte russische Dame, die aus irgendeinem obskuren Grund gehabt hat,

erbat, ich ihren Namen nicht aufzudecken, geschehenen, um mir in Paris zu zeigen, daß das Tagebuch, das sie hatte, die Vergangenheit innen ließ. So ereignislos haben diese Jahre gewesen (angeblich) das das Sammeln täglichen Details (das ist immer eine arme Methode Selbsterhaltung) kaum übertreffen eine kurze Beschreibung vom des Tags Wetter; und es ist neugierig zur Anmerkung respektieren das die persönlichen Tagebücher von Souveränen egal was Mühen überhäufen ihre Reiche machen hauptsächlich bei dem gleichen Thema.

Wer mit Produkten der Maschinenübersetzung vertraut ist, wird vielleicht sogar beeindruckt sein. In der Tat, beide Programme haben brav versucht, ganze Sätze grammatisch zu erfassen, leider falsch, aber immerhin. Unter uns kann man es aber getrost laut sagen: Der Fall ist hoffnungslos. Das ist schlechter als vom schlechtesten Übersetzer. Das ist sogar zu schlecht, um sich redigieren zu lassen, denn an vielen Stellen läßt sich der Sinn nicht einmal erahnen. Zur Erholung darum der gleiche Absatz, vom Humantranslator ins Deutsche gebracht:

Sebastian Knight wurde am 31. Dezember 1899 in der ehemaligen Hauptstadt meiner Heimat geboren. Eine alte russische Dame, die mich aus einem dunklen Grunde bat, ihren Namen nicht zu nennen, zeigte mir in Paris zufällig einmal das Tagebuch, das sie in vergangenen Zeiten geführt hatte. So ereignislos waren jene Jahre (dem Anschein nach) verlaufen, daß die Sammlung täglichen Einerleis (die immer eine armselige Art der Selbstbewahrung ist) kaum mehr enthielt als kurze Wetterbeschreibungen; und es ist merkwürdig, daß sich auch die persönlichen Aufzeichnungen von Staatenlenkern, welche Bedrängnisse ihre Reiche sonst auch heimsuchen, vorzugsweise an denselben Gegenstand halten.

Warum aber ist die Maschinenübersetzung so hoffnungslos schlecht? Warum versagt der Computer so jammervoll? Und wie kann man wissen, daß er auch in der vorhersehbaren Zukunft versagen wird?

Die Antwort darauf findet man, wenn man fragt, was der Computer können müßte, um wenn nicht gut, so doch zumindest richtig zu übersetzen.

Erstens müßte er natürlich alle Wörter im Text kennen. Die heutigen digitalen Lexika sind winzig bis klein bis mittelgroß,

aber sie können ja wachsen. Das Problem aber wäre damit nicht behoben, denn eine natürliche Sprache ist kein geschlossenes, sondern ein offenes System, und in jedem Text können jederzeit Wörter auftauchen, die es noch nie gegeben hat und die niemand voraussehen konnte.

Zweitens brauchte der Computer einen Parser, der jeden Satz bis in seinen letzten Winkel grammatisch zerlegt und auf die in ihm enthaltenen Aussagen reduziert, die dann in der Zielsprache nach deren eigenen syntaktischen Regeln und aus deren Lexikon rekonstruiert werden müßten. Für keine Sprache gibt es bisher einen Parser, der allen ihren erlaubten, also grammatikalischen Satzbauplänen gewachsen wäre. Aber Parser sind immer besser geworden, und was nicht ist, kann ja noch werden.

Das Problem wäre aber auch mit dem optimalen Lexikon und dem idealen Parser keineswegs gelöst, denn natürlich lassen sich Sprachen niemals eins zu eins ineinander zu übersetzen, weder lexikalisch noch syntaktisch. Einmal steht dem Übersetzungsautomaten die Pragmatik jeder Sprache im Weg. Ein Satz mag richtig sein, trotzdem sagt man so einfach nicht. Das spanische *que tal?* wäre mit **what such?* zwar richtig ins Englische übersetzt, bloß dort sagt man etwa *how are you?* Und das hätte dann auf französisch nicht **comment êtes-vous?* zu lauten, sondern *comment ça va?* Auf deutsch wiederum sagt man dafür nicht *wie geht das?*, sondern *wie geht es?*, und wenn man das nun wieder zurück ins Spanische übersetzte, käme jedenfalls nicht *que tal?* dabei heraus.

Ein Stück weit könnte man die Maschine natürlich auf solche pragmatischen Gegebenheiten einrichten. Ihr Hauptproblem besteht aber darin, daß Wörter und Satzphrasen notorisch vieldeutig sind, also in verschiedenen Zusammenhängen Verschiedenes bedeuten und auch unterschiedlich zu übersetzen wären. Das Computerprogramm müßte sich also laufend zwischen verschiedenen Möglichkeiten entscheiden. Es müßte in einem fort desambiguieren. Erst seine kühnen und rührenden Versuche haben uns vorgeführt, wie durch und durch vieldeutig alle Sprache eigentlich ist. Allein mithilfe linguistischer – also lexikalischer, syntaktischer und pragmatischer – Kenntnisse läßt sich keine Sprache desambiguieren. Dazu müßte die Maschine

auch die Semantik beherrschen, also „wissen“, was die Wörter bedeuten. Das heißt, sie müßte über die Definitionen der Begriffe verfügen, wäre aber von vornherein dadurch gehandikapt, daß sich viele Begriffe überhaupt nicht in der erforderlichen Eindeutigkeit definieren lassen, sondern vom menschlichen Geist auf ganz andere Art gebildet werden. Aber selbst die Definitionen würden ihr nicht genügen. Sie müßte außerdem alle interessanten Eigenschaften aller Dinge hinter den Begriffen kennen. Bei dem berühmten Beispielsatz des Philosophen Bar-Hillel *the box is in the pen* müßte sie zum Beispiel wissen, wie groß ein Kasten normalerweise sein kann und wie groß ein *Stift*, um zu entscheiden, daß mit *pen* hier kein Stift gemeint sein kann, sondern vielleicht ein Pferch; aber auch dieses Wissen könnte unbrauchbar sein und etwas ganz anderes, Unvorhersehbares relevant; zum Beispiel die unbegrenzte Fähigkeit des Menschen, auch uneigentliche, also ironische, oder schlicht unsinnige Sätze zu bilden.

In den Zirkeln der Künstlichen Intelligenz spricht man vom sogenannten Weltwissen, das dem Computer beigebracht werden müsse. Er erwirbt es nicht von selbst. Alles muß ihm eindeutig bis ins letzte Detail vorgeschrieben werden. Das heißt praktisch, das dem Programm mit jedem neuen Wort, jeder neuen Wendung nicht nur Dutzende von rein linguistischen Labels, Etiketts angeklebt werden müssen, sondern Labels auch für alles das, was die Menschen üblicherweise von dem betreffenden Ding wissen, obwohl sie gar nicht so recht sagen könnten, was sie von jedem Ding wissen. Jeder neue Eintrag wüchse auf diese Weise, und mancher ins Unermeßliche.

Unter Computerleuten setzt man oft auf die rohe Gewalt, also auf die wachsende Rechenleistung. In der Tat brauchte der Computer eine noch sehr viel größere Rechenleistung, um mit den möglichen Interaktionen all der Labels irgendeines Stückes Sprache fertig zu werden. Aber die bloße Rechenleistung genügte eben nicht. Es müßte ihm erst einmal vorbuchstabiert werden, was er zu berechnen hat. Darum kann eine Übersetzungsmaschine nicht eines Tages erfunden werden. Sie könnte höchstens in jahrzehntelanger Kollektivarbeit Schritt für Schritt aufgebaut werden.

Ich möchte nicht vorhersagen, ob das eines Tages geschehen wird. Vielleicht ja, denn die Menschen sind zäh, wenn sie sich etwas in den Kopf gesetzt haben. Trotzdem würden wir alle nicht arbeitslos. Die Maschine, die es dann gäbe, könnte bestenfalls viele Sätze mehr oder weniger richtig übersetzen. Wir aber wollen nicht nur richtig übersetzen, sondern gut, und das ist etwas ganz anderes.

Manchmal neigen wir dazu, ich selber jedenfalls, unsere Arbeit für ein Handwerk zu halten, eine Art Ingenieurskunst, die nach bestimmten und keineswegs unnennbaren Regeln vonstatten geht. Wir wissen, wo wir nach Wörtern suchen und wie wir Partizipialkonstruktionen auflösen müssen, gut so. Die Sprache im Kopf aber bildet sich nicht als ein kaltes Regelwerk heraus, sondern aus jahrzehntelangen, stark emotional gefärbten persönlichen Spracherfahrungen. Die Wörter und die Satzphrasen hallen in uns wider, gewinnen einen Hof von privaten Assoziationen, nehmen Farbe und Temperatur und Geschmack an, und wir können nur hoffen, nie aber wissen, daß sie das für andere auf die gleiche Weise tun. Diese ganze sensuelle und emotionale Dimension der Sprache aber bleibt dem Computer notwendig verschlossen. Wir könnten nur selber versuchen, einiges davon zu algorithmisieren, kämen aber bald an eine unüberschreitbare Grenze. Sie besteht einfach darin, daß wir uns selber nur höchst unvollständig kennen.

Darum wird es auch noch in hundert Jahren literarische Übersetzer geben, und vielleicht kommen sie auch noch nach Bergneustadt, und einer bekommt den Preis, den Sie heute mir gegeben haben und für den ich hiermit danke.