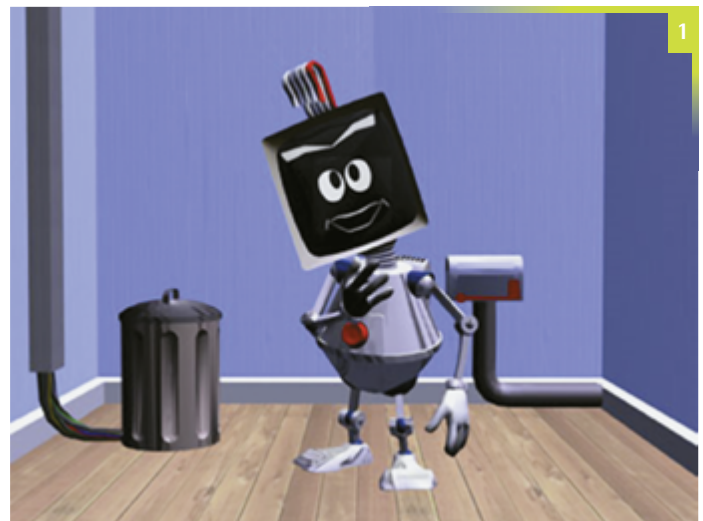


Warum Chatbots keine gekochten Katzen lieben

ARTIFIZIELLE DIALOGAGENTEN UND IHRE SPRACHE

Der Markt für virtuelle Sprachagenten boomt. Spracherkennung, Sprachsynthese und die Implementierung im Netz spielen eine zunehmende Rolle in der Netzkommunikation. Linguistik und Informationswissenschaft können durch die Analyse von Mensch-Mensch-Kommunikation und den statistischen Vergleich mit der Mensch-Maschine-Interaktion einen Beitrag zur Weiterentwicklung der Dialogsysteme leisten.

In dem Zukunftsroman *Otherland* von Tad Williams bildet das Netz eine virtuelle Welt [V(irtual) R(eality)], in der sich Menschen in Form von Simuloiden (sog. Bürger) und Replikanten, Computerprogramme mit dem Aussehen eines Menschen, bewegen und miteinander kommunizieren. Die sprachliche Kommunikation zwischen Bürger und Replikant erfolgt reibungslos und ohne Probleme. In dem Roman wird das Verhältnis von Realität und Virtualität (Cyberspace) zu einem Schreckensszenario verdichtet und das Computernetzwerk als eine Art neue Spezies gesehen, ein *Apparatus sapiens*, der sich koevolutionär mit dem *Homo technicus* entwickelt hat. Sprach- und Denkfähigkeit, geistige Kompetenz der Maschine werden als Resultat dieser Evolution begriffen. Nun sind wir glücklicherweise von einer solchen Entwicklung weit entfernt, aber: die Schnittstellen zwischen Mensch und Technosphäre werden immer enger, reale Welt und virtuelle Welt beginnen sich zu überlappen. In Bezug auf die sprachliche Kommunikation lässt sich in den letzten Jahren beobachten, dass Chatprogramme und Lingubots, also künstliche Gesprächspartner, in der Netzkommunikation deutlich zugenommen sowie ihre Leistungsfähigkeiten erheblich verbessert haben. Spracherkennung, Sprachsynthese und die Implementierung im Netz



spielen zunehmend eine Rolle in der Netzkommunikation, aber auch in der Interaktion des Users mit Handy, Netbook etc. Wir wollen im Folgenden auf die Schnittstelle der Mensch-Maschine-Interaktion (human-computer interaction, HCI) im Hinblick auf sprachliche Dialoge eingehen und insbesondere auf User-Sprachagenten-Interaktion, ein Forschungsschwerpunkt, der uns im Rahmen des Projekts [mediensprache@net](#) grundsätzlich und anwendungsbezogen (System *El Lingo*) interessiert.

Chatbots und ECAs

Mit artifiziellen Dialogagenten können wir natürlichsprachlich in Interaktion treten. Auf kommerziellen Homepages im

Netz agieren sie als virtuelle Kundenberater, die über die unterschiedlichsten Themen informieren. In E-Learning-Anwendungen fungieren sie als künstliche Lehrer, die Fragen stellen können und Feedback geben. Im Entertainment-Bereich trifft man sie als unterhaltsame Gesprächspartner mit Small-Talk-Funktion zum Plaudern an. Darüber hinaus moderieren Bots Chatrooms und treten als artifizielle Charaktere in *MMORPGs* (Massively Multiplayer Online Role-Plays, interaktive Spiele im Internet, z.B. »World of Warcraft«) oder VR-Umgebungen auf.

In Deutschland boomt der Markt für virtuelle Sprachagenten. Wenn Sie sich zum Beispiel für ein Auslandsstudium interessieren, beantwortet

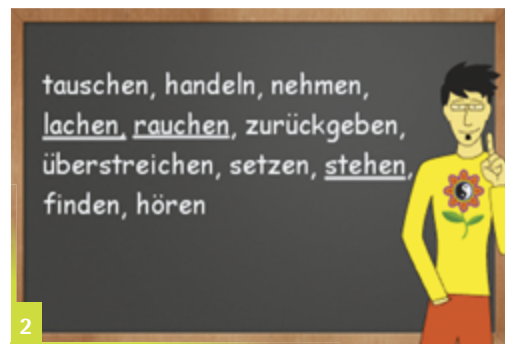
Ihnen *Luzie*, die künstliche Studienberaterin des DAADs, alle bereichsspezifischen Fragen (<http://bot.daad.de/>), während der virtuelle Bundesadler *Findulin* auf <http://adler.bundestag.de/> Wissenswertes zur deutschen Verfassung ausgibt. International und multilingual im Einsatz ist die charmante Kundendienstmitarbeiterin *Anna* von Ikea (<http://www.ikea.com/de/>). Gewinner des Loebner-Preises 2008 ist der übellaunige Roboter *Elbot* der Hamburger Firma Artificial Solutions, der seine User mit sarkastischen Sprüchen unterhält.

IT-Firmen vertreiben einfach angelegte Chatbotssoftware, mit der jeder Betrieb und jede öffentliche Einrichtung den eigenen Bot programmieren und ins Netz stellen können. Denn hinter den klassischen Chatbots steckt oft nicht mehr als eine simple Schlüsselworterkennung. Die neue Generation von Chatbots sind dagegen so genannte ECAs – *Embodied Conversational Agents*. Sie sind nach menschlichem Vorbild gestaltet und interagieren multimodal sowie zielorientiert. Hinter oft aufwendig animierten Avataren und individuell designten Agenten-Personae verbergen sich Systemarchitekturen, die nicht nur auf umfangreiche, semantische Lexika zurückgreifen können, sondern auch Dialogwissen miteinbeziehen. Parser zur Analyse der User-Eingaben greifen auf große Datenbanken und elaborierte Wahrscheinlichkeitskalküle zurück. Viele ECAs verfügen über Sprechcharaktererkennung, einfache Syntaxparser, Modelle zum Turn-Taking etc.; darüber hinaus können dynamische Usermodelle aufgebaut werden, die es dem Bot ermöglichen, sich an seine Interaktionspartner zu »erinnern«.

Auch am Deutschen Seminar der Leibniz Universität Hannover wird ein Chatbot-System

für den Einsatz als virtueller Tutor in einer E-Learning-Umgebung entwickelt. Das System *El Lingo* verfügt über eine einfache, aber sehr robuste Architektur und ein klares Dialog-Design. Der freundliche Avatar soll Studierende der Germanistik als feedbackgebender Helfer durch eine Online-Einführung in die Sprachwissenschaft führen zur Unterstützung der curricularen Lehre und des Selbststudiums. Die benutzerfreundliche Software zur Gestaltung der Dialoge kann auch auf die Belange anderer Fachbereiche abgestimmt werden.

verschiedlichen Erwartungen. Außerdem bestehen im Dialog mit dem künstlichen Gegenüber auf dem heutigen Entwicklungsstand Möglichkeiten, die sich im Gespräch unter Menschen nicht bieten, da sozial zu viel auf dem Spiel stünde. Im Klartext heißt das, auch die besten Bots werden immer noch beschimpft, einsilbig herumkommandiert und permanent auf ihre Parsing-Fähigkeiten hin getestet. In den Untersuchungskorpora finden sich große Mengen von Belegen, die die Vermutung nahe legen könnten, es existiere eine Art *Computer-Talk* als sprach-

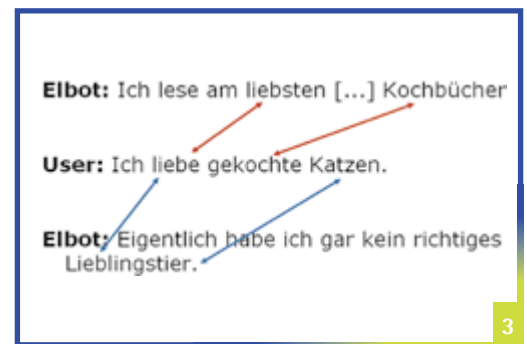


Herausforderungen an die Linguistik

Aus linguistischer Perspektive interessieren zwei übergeordnete Fragestellungen:

1. Welche Aspekte menschlichen Dialogverhaltens sind antizipierbar und lassen sich somit für ein deterministisches System programmieren?
2. Wie muss ein Dialog-Design konkret umgesetzt werden, um sich diese Aspekte menschlichen Verhaltens zu Nutze zu machen?

Die systematische Korpusanalyse von Dialogtranskripten unterschiedlicher Bots zeigt, dass der Dialog mit artifiziellen Agenten von verschiedenen Variablen beeinflusst wird. So begegnen Menschen dem System mit unterschiedlichen Vorannahmen, unterschiedlicher Technikexpertise und un-



liches Register, wie Magdalena Zoeppritz sie schon 1988 annahm. Dagegen steht die Beobachtung, dass sich trotz teilweise notorisch schwieriger *Verstehensprobleme* auf Seiten des Systems bemerkenswerte Anpassungsprozesse seitens der User an die Systemsprache vollziehen. Letzteres wird im linguistischen Fachdiskurs als *Alignment* bezeichnet und als ein Mechanismus verstanden, der Interaktion dynamisch regelt. Menschen tendieren generell dazu, sprachliche Strukturen ihres Gegenübers zu spiegeln. Dies geschieht sowohl auf lexikalischer Ebene durch wörtlichen Bezug und Sinnrelationen als auch auf syntaktischer Ebene zum Beispiel durch Parallelkonstruktionen:

Einem solchen Anpassungsverhalten liegt nach PICKERING UND GARROD (2004) kognitiv ein

Abbildung 1
Elbot, anthropomorphes Design mal anders
Quelle: Fred Roberts, *Artificial Solutions*, 2008

Abbildung 2
El Lingo erklärt das Konzept der Verbovalenz
Quelle: Torsten Siever, Netaya Lotze, *LUH*, 2011

Abbildung 3
Elbot-Korpus 3590–3627



Prof. Dr. Peter Schlobinski

Jahrgang 1954, ist seit dem Wintersemester 1995/96 Professor für Linguistik am Deutschen Seminar. Kontakt: pschlobi@uos.de



Netaya Lotze

Jahrgang 1979, ist seit 2008 wissenschaftliche Mitarbeiterin im Bereich Linguistik am Deutschen Seminar. Kontakt: netayalotze@web.de

Lower-Level-Priming-Mechanismus zu Grunde, was bedeutet, dass Dialoge bis zu einem gewissen Grad ohne Einbindung von höheren Kognitionsprozessen geführt werden können. Ein solch mechanistischer Zugang zur dialogischen Interaktion kann für HCI-Anwendungen insofern fruchtbar gemacht werden, als mechanistische Userreaktionen von einem deterministischen System unter Idealbedingungen zuverlässig ausgelöst werden können müssten. Erkenntnisse über ein regelhaftes, verlässliches Userverhalten – egal ob in Bezug auf *Computer-Talk* als allgemeines Register, in Abhängigkeit vom Usertyp oder als dynamischer Alignment-Mechanismus – sind hilfreich für die Weiterentwicklung der Systeme. Doch trotz allem lässt sich menschliches Dialogverhalten nur sehr bedingt vorhersagen, da Menschen eben nicht wie Sprachsysteme rein deterministisch agieren. Eine spontane Äußerung wie *Ich liebe gekochte Katzen.* im Kontext des obigen Dialogs kann zwar wunderbar als Beleg für Alignment angeführt werden, ließe sich aber kaum antizipieren, weshalb sie für den Systemparser eine so komplexe Aufgabe darstellt. Hier zeigen sich deutlich die Grenzen der aktuellen Genera-

tion von Sprachverarbeitungssystemen. Semantisches Wissen ist in den meisten Systemen nur rudimentär angelegt, eine kognitive Komponente im Sinne logischen Schließens fehlt bislang ganz. So sind Bots nicht in der Lage, inferenzielle Schlüsse zu ziehen:

- 1 USER: bist du ein mensch
- 2 SYSTEM: nein aber ich bemühe mich menschlich zu sein
- 3 USER: du bist aber dumm
- 4 SYSTEM: Meine Mama hat immer gesagt, dumm ist der der Dummes tut.
- 5 USER: ja, also du bist dumm, denn ich bin deine mutter!
- 6 SYSTEM: Meine Mama hat immer gesagt, dumm ist der der Dummes tut. Wie alt bist du?
(Max-Korpus 93489–93494)

In dem obigen Beispiel wiederholt das System als Ausweichstrategie für Nichtverstehen in T. 6 die Äußerung aus T. 4 und initiiert einen Themawechsel. Eine menschliche Reaktion wäre entweder (a) die epistemische Begründung in T. 4 (*Du bist nicht meine Mutter.*) oder (b) die logische fehlerhafte Argumentation zurückzuweisen. Das System wird gemäß der Aussage, *Dumm ist der, der dummes tut.* nicht dadurch als *dumm* charakterisiert, dass dies von der Mutter behauptet wird, sondern lediglich durch das Ausführen *dummer* Handlungen. Äußerung 5 wären nur dann richtig, wenn in Äußerung 4 gesagt worden wäre: *Wenn meine Mutter sagt, dass ich dumm bin, dann bin ich dumm.* Dann folgte mit Äußerung 3 und der Feststellung in 5, dass er die Mutter des Systems sei, und modus ponens, dass das System dumm sei. Das System ist eben nicht programmiert, aus den sprachlichen Daten argumentative Schlüsse zu ziehen und komplexe Argumentationsschemata zu erstellen.

Die Zahl der antizipierbaren und manipulierbaren Aspekte von Dialogen ist gering: ritualisierte Strukturen wie Begrüßungen und Verabschiedungen, bestimmte Adjazenzphänomene und mit einer gewissen Wahrscheinlichkeit Alignment sowie konventionalisierte semantische Assoziationen – alles in Abhängigkeit von den spezifischen Variablen des jeweiligen Users und des jeweiligen Systems. Linguistik und Informationswissenschaft können durch die Analyse von Mensch-Mensch-Kommunikation und dem statistischen Vergleich mit der Mensch-Maschine-Interaktion einen Beitrag zur Weiterentwicklung der Dialogsysteme leisten: Explorativ in Korpusstudien, Experimenten und Usability-Tests, modellgestützt auf der Grundlage psychologischer oder psycholinguistischer Modelle oder theoriegeleitet zum Beispiel auf Grundlage semantischer Netze zur Entwicklung eines möglichst menschenähnlichen Recipient-Designs.

Literatur

- Buschmeier, H., Bergmann, K., & Kopp, S. (2010): Modelling and Evaluation of Lexical and Syntactic Alignment with a Priming-Based Microplanner. In E. Kraemer, & M. Theune, Empirical methods in natural language generation. (pp. 85–104). Berlin/Heidelberg: Springer.
- Williams, Tad (1998 f.). Otherland, Vol. I–IV. Stuttgart.
- Zoeppritz, Magdalena (1989). Computer talk? Technical Report TN 85.05, IBM Heidelberg Scientific Centre.