



NETWORX Nr. 25

Michaela Storp

**Chatbots. Möglichkeiten und Grenzen
der maschinellen Verarbeitung
natürlicher Sprache.**

(2002)

online verfügbar seit: 2002-02-02

URL: <http://www.websprache.net/networx/docs/networx-25.pdf>

Versionshinweise: keine



NETWORX – Online-Publikationen zum Thema Sprache und Kommunikation im Internet:
<http://www.websprache.net/networx/>

Inhaltsverzeichnis

1 Einleitung.....	3
2 Turing Test und Loebner Contest	5
2.1 Can machines think? Der Turing Test.....	5
2.2 Der Loebner Contest.....	7
3 Die Anfänge - ELIZA.....	9
4 Von ELIZA zu ALICE.....	12
5. Intelligente Sprachverarbeitung?	15
5.1 Vom Verstehen natürlicher Sprache	16
5.2 Die Funktionsweise von Chatbots - ALICE	19
5.3 Tricks	26
6 Fazit	29
Literatur	32
Links (Auswahl)	33

1 Einleitung

1950 prophezeite der englische Mathematiker Alan Turing, in fünfzig Jahren werde es möglich sein, Computer darauf zu programmieren, menschliches Gesprächsverhalten so erfolgreich zu imitieren, dass es kaum mehr möglich sein werde, Computer und menschlichen Gesprächspartner zu unterscheiden.¹

Seit dieser Prophezeiung sind inzwischen gut fünfzig Jahre vergangen. Heute ist es tatsächlich möglich, im Internet Gespräche mit Computern zu führen. Auch der folgende Dialog wurde nicht zwischen zwei Menschen, sondern zwischen Mensch und Maschine geführt:

- Hi John, wie geht's?
- Wirklich? Wir fühlen uns sehr geehrt.
- Die Ehre ist ganz auf unserer Seite.
- Letztendlich macht es keinen Unterschied.
- Wir sind so froh, dass Sie noch am Leben sind.
- Das ist nicht wirklich wichtig.
- Aber Sie sind doch noch am Leben?
- Huh, wer kann das schon sagen?
- Also ja oder nein?
- Ich muss mit Ringo sprechen und sehen, wie er dazu steht.
- Wissen Sie, das ist auch der Grund warum es mir immer um Frieden ging.

Dialoge wie dieser kurze Chat mit „John Lennon“ sind heute an vielen Stellen im Internet möglich. Programme wie „John Lennon“ werden Chatbots, Chatterbots oder auch Chat-Roboter genannt. Sie zielen darauf, einen menschlichen Gesprächspartner zu simulieren. Dabei wird versucht, den menschlichen Umgang mit natürlicher Sprache so geschickt nachzuahmen, dass es zumindest zeitweilig gelingt, den menschlichen Gesprächsteilnehmer zu täuschen.²

¹ vgl. Turing (1950)

² Laven: „A chatterbot is a program that attempts to simulate typed conversation, with the aim of at least temporarily fooling a human into thinking they were talking to another person.“ www.simonlaven.com

Zur Zeit erleben diese virtuellen Gesprächspartner einen wahren Boom. Vetter schätzt, dass wöchentlich ein neuer Chatbot ins Internet gestellt wird.³ Neben dem reinen Chat übernehmen die Bots auch andere Aufgaben: Verwaltung von IRC-Foren, Informationssuche im WWW, Internet-Auktionen, Kundenservice in Online-Shops, künstliche Charaktere in MUDs usw.

Die vorliegende Arbeit widmet sich den Chatbots im engeren Sinn, die ausschließlich für Online-Gespräche mit Menschen programmiert wurden. Dabei kann zunächst zwischen kommerziellen und nicht-kommerziellen Varianten unterschieden werden.

Immer mehr Firmen ergänzen ihren Internetauftritt mit einem Chatbot, der als virtueller Gastgeber fungiert, durch die Website führt und Fragen zu Produkten und Dienstleistungen beantwortet. Davon erhofft man sich eine stärkere Kundenbindung, Einsparungen im Bereich der Telefon-Hotlines und Call-Center und nicht zuletzt - durch Auswertung der Gesprächsprotokolle - Marketingformationen. Entsprechend gibt es kommerzielle Anbieter von Chatbot-Editoren, die es dem Endanwender vereinfachen, einen eigenen Chatbot zu programmieren.

Es existiert außerdem eine Reihe nicht-kommerzieller Chatbots. Sie werden in erster Linie aus Spaß am Experimentieren entwickelt, nicht zuletzt in der Hoffnung, ein wirklich intelligentes Programm zu schaffen. Es sind überwiegend die Entwickler dieser nicht-kommerziellen Chatbots, die sich jährlich zum Wettbewerb um den Loebner Preis treffen, bei dem das menschenähnlichste Programm in einem abgewandelten Turing Test ermittelt wird.

Diese Arbeit soll einen ersten Einblick in die Welt der Chatbots geben. Dabei geht es weniger um programmiertechnische Details als um die Frage der theoretischen Machbarkeit. Werden die heutigen Chatbots dem Anspruch ihrer Entwickler nach intelligenter Sprachverarbeitung gerecht? Können Computer überhaupt natürliche Sprache verstehen? Und welche Schwierigkeiten treten dabei auf? Sind Chatbots ein Beispiel für „intelligente“ Sprachverarbeitung oder handelt es sich um „Intelligenzbetrugsprogramme“?

³ Vetter 2001:228

2 Turing Test und Loebner Contest

Seit 1991 treffen die Entwickler von Chatbots jährlich zum Wettbewerb um den Loebner Prize zusammen. In diesem Wettbewerb konkurrieren die Chatbots um die Auszeichnung als „mensenähnlichstes“ Programm. Fernziel des Loebner Contest ist es jedoch, den nach dem gleichnamigen englischen Mathematiker benannten Turing Test zu bestehen. Damit wäre eine Unterscheidbarkeit von Chatbot und Mensch als Gesprächspartner nicht mehr gegeben. Dem Chatbot wäre somit menschenähnliches (intelligentes?) Sprachverhalten zu unterstellen. Worum geht es bei diesem in der Chatbot-Szene als Maßstab geltenden Turing Test? Wir geben im Folgenden einen Einblick in diesen noch immer kontrovers diskutierten Gegenstand⁴.

2.1 Can machines think? Der Turing Test

Können Maschinen denken?

Mit dieser Frage leitete Alan M. Turing, britischer Mathematiker, seinen 1950 in der Zeitschrift *Mind* erschienenen Aufsatz „Computing machinery and intelligence“ ein. Aufgrund des Problems einer klaren Definition sowohl von *Maschine* als auch von *denken*, ersetzt Turing die Frage nach der Denkfähigkeit von Maschinen durch ein Gedankenexperiment, das so genannte Imitationsspiel.

In seiner ursprünglichen Form wird das Imitationsspiel mit drei Personen durchgeführt: Einem Mann (A), einer Frau (B) und einem Interviewer beliebigen Geschlechts (C). Der Interviewer ist von den beiden anderen Personen räumlich getrennt; der Kontakt erfolgt über maschinengeschriebene Sprache, um Hinweise durch Stimme, Handschrift etc. auszuschließen. Der Interviewer hat die Aufgabe zu bestimmen, wer von den Mitspielern die Frau ist. Dazu stellt er den beiden Fragen. Die Schwierigkeit besteht darin, dass beide Personen behaupten, weiblich zu sein. Der Mann täuscht also vor, die Frau zu sein, er imitiert eine Frau.

⁴ Eine detaillierte Schilderung der Diskussion um den Turing Test würde den Schwerpunkt dieser Arbeit verschieben. Wir verweisen daher auf Turing (1950) sowie auf Saygin, Cicekli, Akman (2000).

In einem weiteren Gedankenschritt benutzt Turing das Imitationsspiel als Methode zur Klärung der Eingangsfrage. Was würde passieren, wenn eine Maschine den Part von A übernehme? Würde der Interviewer genauso oft bei der Zuordnung irren? Kann also die Maschine genauso gut eine Frau imitieren wie der Mann? Schließlich beschränkt Turing den Begriff *Maschine* auf *Computer* und schlägt eine weitere Variante des Imitationsspiels vor: Der Computer übernimmt die Rolle von A, ein Mann die Rolle von B. Beide geben vor, weiblich zu sein.

„Let us fix our attention to one particular digital computer *C*. Is it true that by modifying this computer to have an adequate storage, suitably increasing its speed of action and providing it with an appropriate programme, *C* can be made to play satisfactorily the part of *A* in the imitation game, the part of *B* being taken by a man.“⁵

Turing war überzeugt, Computer könnten, entsprechend große Speicherkapazität und Rechengeschwindigkeit vorausgesetzt, so programmiert werden, dass sie das Imitationsspiel erfolgreich spielen. Bei Turing ersetzt die Frage nach der Kommunikationsfähigkeit schließlich die eingangs gestellte Frage nach der Denkfähigkeit. Gelänge es einem Computer(programm) tatsächlich, die Kommunikationsfähigkeit eines Menschen in natürlicher Sprache zu imitieren, so müsste man der Maschine Intelligenz unterstellen. Vorausgesetzt, dem Computer gelänge es dauerhaft und in unabhängigen Versuchen zu täuschen.⁶ Die Fähigkeit mit natürlicher Sprache umzugehen ist eng verknüpft mit Denken und Intelligenz. Und somit wird das Imitationsspiel schließlich zum (Intelligenz-)Test, dem Turing Test (TT).

Es liegt nahe, dass Turings Einschätzung seines Tests zu Widerspruch aus den unterschiedlichsten Richtungen führte. Die Diskussion um den TT ist seit der Veröffentlichung 1950 nicht abgerissen. Während einige in ihm den Ausgangspunkt der künstlichen Intelligenz (KI) sehen, halten andere ihn für

⁵ Turing (1950:442) Es bleibt unklar, ob der Computer mit einer Frau oder einem Mann spielen soll. Heute wird Turings Ansatz dahingehend verstanden, dass unabhängig vom Geschlecht das Vermögen eines Computers, einen Menschen zu imitieren, ermittelt werden soll. vgl. Saygin, Cicekli, Akman (2000:4).

⁶ Saygin, Cicekli, Akman (2000:35): „Turing required *consistently* successful performance from machines to grant them intelligence.“

irreführend und als Intelligenztest für völlig unzureichend. Turing selbst hat in seinem Aufsatz zahlreiche Einwände religiöser, philosophischer, mathematischer oder biologisch-psychologischer Art gegen seine Idee von der denkenden Maschine antizipiert und zu widerlegen versucht.⁷

Während es in der KI-Forschung auch heute, 50 Jahre nach Turing, keine Einigung über die Aussagekraft des TT gibt, scheint er den Chatbot-Entwicklern als maßgebliche Orientierung zu dienen. Ihr Ziel ist es, ein System zu entwickeln, das den TT besteht. Der Chatbot soll im „Gespräch“ nicht mehr vom Menschen zu unterscheiden sein; einen menschlichen Gesprächspartner perfekt imitieren.

2.2 Der Loebner Contest

Der TT wurde in seiner von Turing ursprünglich vorgesehenen Form nie realisiert. Die Diskussion um die Denkfähigkeit von Maschinen blieb vor allem im Bereich der Sprachbeherrschung lange der Theorie verhaftet. Erst als gegen Ende der 80er Jahre ausreichend leistungsfähige Computer entwickelt wurden, konnte mit praktischen Versuchen ernsthaft begonnen werden.

Seit der Amerikaner Hugh Loebner 1990 die *Loebner Prize Competition in Artificial Intelligence* ins Leben rief, wird jährlich ein modifizierter TT durchgeführt, dem sich verschiedene Chatbot-Programme stellen. Der mit 100.000 US-Dollar und einer Goldmedaille dotierte Preis geht an das Programm, dem es gelingt, den Test zu bestehen.⁸ In jedem Jahr werden außerdem eine Bronzemedaille und 2.000 US-Dollar für das „mensenähnlichste“ Programm vergeben.

Der erste Loebner Contest fand 1991 im *Boston Computer Museum* statt. Er war zugleich die erste Durchführung eines TT. Sechs Programmen und vier Menschen standen zehn Interviewer gegenüber. Da man von vornherein davon ausging, dass es keinem der Programme gelingen würde, den TT zu bestehen,

⁷ Die bereits von Turing diskutierten Einwände, sowie weitere Gegenargumente werden ausführlich in Saygin, Cicekli, Akman (2000) diskutiert.

⁸ Loebner fordert seit 2000 zudem die Beherrschung von audio-visuellem Input (Saygin, Cicekli, Akman 2000:35) Einzelheiten zu den Regeln des Loebner Prize unter: www.loebner.net/Prize/loebner-prize.html

wurden die Gespräche inhaltlich auf ein vorgegebenes Thema beschränkt. Dennoch fiel es den Juroren nicht schwer, die Programme von den Menschen zu unterscheiden.⁹ Turing selbst hatte zu einer so stark eingeschränkten Testvariante angemerkt, dass sie keine Schlüsse auf die Intelligenz eines Systems zuließe.

Seit 1995 ist die thematische Restriktion aufgehoben. Die Länge der einzelnen Gespräche bleibt dem Ermessen der Juroren überlassen. Jeder von ihnen hat zudem die Möglichkeit, sich mehrfach mit den einzelnen Programmen zu befassen. Die Hauptaufgabe der Interviewer besteht darin, die teilnehmenden Systeme entsprechend ihrer „Menschlichkeit“ in eine Rangfolge zu bringen. Außerdem gilt es, den Punkt zu bestimmen, an dem ein Programm vom Menschen nicht mehr zu unterscheiden ist.

Bisher ist es keinem Programm gelungen, diese Variante des TT zu bestehen. Die von Loebner neu hinzugefügte Bedingung der Verarbeitung von audiovisuellem Input und entsprechendem Output, stellt die Entwickler vor zusätzliche Schwierigkeiten. 1999 wurde ein weiteres Etappenziel eingeführt. Demnach erhält ein Wettbewerber eine Silbermedaille und 25.000 US-Dollar, wenn sein Programm eine „Irreführungsquote“ von 50% erreicht. Auch dies ist bislang nicht erreicht worden.

Der Loebner Contest ist nicht weniger umstritten als der TT selbst. Loebner selbst sieht die Förderung der KI-Forschung als Ziel seines Wettbewerbs; die Mehrheit der Wissenschaftler steht ihm jedoch ausgesprochen kritisch gegenüber.¹⁰ Während einige Stimmen aus der KI es zumindest für möglich halten, dass sich ein Loebner-Teilnehmer zu einem nützlichen natürlichsprachigen System entwickeln könnte, lehnen andere den Wettbewerb als fehlgeleitet und falsch verstandene Forschung ab.¹¹

⁹ Auch wenn, so eine oft zitierte Anekdote, eine menschliche Teilnehmerin aufgrund ihres ausgesprochen detaillierten Wissens über Shakespeare irrtümlich für einen Computer gehalten wurde. s. u.a. Saygin, Cicekli, Akman 2000:36

¹⁰ Mit Hutchens (1997) kritisiert selbst der Gewinner der Bronzemedaille von 1997 den Wettbewerb als unwissenschaftlich und der seriösen KI-Forschung nicht zuträglich. Hutchens 1997:19

¹¹ Saygin, Cicekli, Akman 2000:36; Als Beispiel für die Polemik um den Loebner Prize mag Marvin Minsky, etablierter KI-Forscher, dienen. Er hat demjenigen \$100 versprochen, dem es gelingt, Loebner zur Aufgabe seines Wettbewerbs zu bringen. Loebner konterte, indem er Minsky als Co-Sponsor seines Preises betitelte. Denn, so lauten die Regeln, der Wettbewerb wird eingestellt, sobald es einem Kandidaten gelingt, den TT zu bestehen. Das Preisgeld stiege, dank Minsky, auf \$100.100.

Obwohl bisher nur Bronzemedailles vergeben wurden, ist die Leistungsfähigkeit der teilnehmenden Programme seit 1991 erheblich verbessert worden.¹² Wir werden in Abschnitt 4 auf die bisherigen Gewinner eingehen.

3 Die Anfänge - ELIZA

Turings Gedankenexperiment und die daraus entwickelte Testsituation dienen den heutigen Chatbot-Entwicklern als Maßstab. Bevor wir uns jedoch mit den aktuellen Programmen und ihrer Leistungsfähigkeit befassen, sollen die programmier-technischen Anfänge sprachbegabter Programme vorgestellt werden. In den Jahren 1964-66 entwickelte Joseph Weizenbaum am MIT ein Programm, das erstmals die Mensch-Maschine-Kommunikation in natürlicher Sprache ermöglichte: ELIZA. Und obwohl Weizenbaum mit diesem System nicht auf den Turing Test zielte, gilt ELIZA heute als der erste Chatbot.

ELIZA simuliert einen Psychotherapeuten in der Gesprächstherapie nach Rogers. Weizenbaum wählte diese Rolle, da er sie für einfach umzusetzen hielt: Die Aktivität des Therapeuten beruht nämlich im Wesentlichen darauf, die Aussagen des Patienten als Frage zu paraphrasieren, sie zu „spiegeln“ und so den Patienten zum Weiterreden zu motivieren.¹³ Der Therapeut verzichtet auf vollkommen auf eigene Gesprächsinitiativen, reagiert nur auf Gesagtes. Dementsprechend verhält sich ELIZA, wie folgendes Transkript verdeutlicht¹⁴:

- I need some help, that much seems certain.
- WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP
- Perhaps I could learn to get along with my mother
- TELL ME MORE ABOUT YOUR FAMILY
- My mother takes care for me
- WHO ELSE IN YOUR FAMILY TAKES CARE FOR YOU

¹² An dieser Stelle sei auf die Transkripte der Gespräche mit den jeweiligen Wettbewerbssiegern auf der Loebner Homepage verwiesen.

¹³ Weizenbaum 1976:3; ELIZA konnte auch auf andere Rollen (Skripts) programmiert werden. Sie wurde jedoch in dieser, so genannten Doktor-Version bekannt.

¹⁴ Weizenbaum 1976:4 / Antworten von ELIZA in Großbuchstaben

- My father
- YOUR FATHER
- You are like my father in some ways
- WHAT RESSEMBLANCE DO YOU SEE
- [...]
- You don't argue with me
- WHAT MAKE YOU THINK I DON'T ARGUE WITH YOU
- You are afraid of me
- DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

ELIZA ist ein einfaches Reiz-Reaktions-System, das aus zwei Hauptkomponenten besteht¹⁵: Das eigentliche Programm steuert das Einlesen der Nutzereingabe, die Ausgabe der Antworten auf dem Bildschirm sowie die Einbindung des Skripts. Dieses Skript ist die zweite Komponente und könnte als das Gehirn ELIZAs bezeichnet werden. Es enthält die Erkennungsmuster für die Analyse der Nutzereingaben sowie entsprechende Antworten.

Die Eingabe des Nutzers, die damals noch über eine an den Computer angeschlossene Schreibmaschine erfolgte, wird zunächst auf Schlüsselwörter untersucht. Außerdem bewirken so genannte Ersatzschlüssel den Austausch der Personalpronomen der ersten und dritten Person. Der nächste Schritt ist ein Kontextmuster-Test. Dabei wird die Eingabe mit einer gespeicherten Liste von (syntaktischen) Mustern für das jeweilige Schlüsselwort verglichen. Findet sich kein passendes Muster, wird eine Auffangantwort generiert („I see.“). Zu jedem Kontextmuster existiert eine Reihe von Antwort-Rümpfen, aus denen per Zufallsgenerator ausgewählt wird. Die Rümpfe enthalten Leerstellen, die entsprechend der syntaktischen Kontextanalyse aufgefüllt werden. Die so erzeugte Antwort wird auf dem Bildschirm ausgegeben. ELIZA arbeitet rein syntaktisch. Die Suche nach Schlüsselwörtern ist als einfacher Zeichenabgleich zu verstehen. Wörter sind für ELIZA Zeichenketten ohne Bedeutung, die in bestimmten syntaktischen Kontexten auftreten können.

¹⁵ Eine detaillierte Beschreibung der Funktionsweise sowie Beispiel-Skripte von ELIZA findet sich bei Steup 1995:171ff

ELIZA wurde 1966 euphorisch aufgenommen. Weizenbaum hingegen war von den seiner Einschätzung nach übertriebenen Reaktionen schockiert:

„I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. [...] What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people.“¹⁶

Weizenbaum hielt zudem sein Programm für überbewertet. Mehrere Psychologen sahen in ELIZA die Geburtsstunde der automatisierten Therapie. Eine Vorstellung, die auch heute trotz des selbstverständlichen Umgangs mit Computern noch erschreckt. Und während Weizenbaum in ELIZA vor allem den Beweis der Kontextabhängigkeit (im Sinne von syntaktischem Kontext) von Sprachverstehen sah¹⁷, wurde das Programm in KI-Kreisen als Durchbruch beim maschinellen Verstehen natürlicher Sprache betrachtet. Eine Einschätzung die schon bald durch die weitere KI-Forschung revidiert wurde. Es zeigte sich, dass die Verarbeitung natürlicher Sprache ein weitaus komplexeres Problem darstellt, das sich nicht allein durch syntaktische Analysen lösen lässt und dem man sich in Teilschritten nähern muss.

ELIZA aber existiert auch heute noch in zahlreichen Versionen im Internet und kann sich auch nach fast 30 Jahren noch mit einigen der modernen Chatbots messen. So gewann z. B. eine stark an ELIZA angelehnte Lösung drei Mal den Loebner Prize in dessen Anfangsjahren.¹⁸ Viele moderne Chatbots können als direkte Weiterentwicklungen ELIZAs betrachtet werden.

¹⁶ Weizenbaum 1976:6f

¹⁷ Weizenbaum 1976:7

¹⁸ Weintraubs System PC Therapist wurde 1991-93 und 1995 die Bronzemedaille verliehen. Es gilt als ELIZA-like.

4 Von ELIZA zu ALICE

Obwohl Weizenbaum bei der Entwicklung von ELIZA keineswegs den Turing Test anvisierte und 1966 an Chatbots im heutigen Sinn noch nicht zu denken war, gilt ELIZA rückblickend als „Mutter der Chatbots“. Tatsächlich war ELIZA das erste Programm überhaupt, das Mensch-Maschine-Kommunikation in natürlicher Sprache ermöglichte. Aus heutiger Sicht erscheinen die Fähigkeiten dieses Systems zwar eher bescheiden, doch begann mit ELIZA die experimentelle Phase der maschinellen Verarbeitung natürlicher Sprache. Die Frage nach der Kommunikationsfähigkeit von Computern blieb nicht länger der Theorie überlassen.

Im Anschluss an ELIZA widmete sich die junge Disziplin der KI im Bereich der Sprachverarbeitung überwiegend Teilzielen. Chatbots entstanden nur vereinzelt. Dieser Forschungsbereich scheint in der Tat erst durch Loebners Ausschreibung motiviert worden zu sein. Die Entwicklung und Verbreitung des Internets erlaubte es schließlich, die Systeme aus den Laboren heraus einer breiten Öffentlichkeit vorzustellen und an ihr zu testen.

In den 70er Jahren programmierte Kenneth Colby PARRY, ein System, das einen Paranoiker simuliert. Dabei strebte Colby ebenso wenig den TT an wie Weizenbaum. Sein Ziel war es, durch die Computersimulation eine Theorie über typisch paranoides Kommunikationsverhalten zu bestätigen.¹⁹ In einem veränderten TT ließ Colby dennoch sein Programm gegen menschliche Paranoiker antreten. Dabei hatten selbst erfahrene Therapeuten Schwierigkeiten, Computer und Mensch zu unterscheiden. Genau wie ELIZA ist PARRY jedoch nur in einem sehr eingeschränkten Bereich fähig, einen Menschen zu simulieren. Dabei macht sich das Programm die Eigenheiten psychisch kranker Menschen zunutze. Eventuelle Schwächen oder Fehler im sprachlichen Bereich werden vom Therapeuten schnell als Symptome der psychischen Störung begriffen, nicht aber als Fehler der maschinellen Sprachverarbeitung.²⁰ PARRY verzichtet auf ein syntaktisches Parsing oder Desambiguierungsmechanismen. Das Programm

¹⁹ vgl. Hutchens (1997:5)

²⁰ Saygin, Cicekli, Akman (2000:34)

sucht lediglich nach Schlüsselwörtern in der Eingabe. Ein so einfaches Verfahren kann nur in einem derart eingeschränkten Rahmen erfolgreich sein. Aber eben in genau dieser Kommunikationssituation (paranoider Patient spricht mit Therapeuten) entspricht die Programmierung genau den Anforderungen, denn es stellt, so Saygin, Cicekli, Akman, eine gelungene Annäherung an die typischen kognitiven Prozesse des Paranoiden dar.²¹

Gegen Anfang der 90er Jahre begann die Zeit der Chatbots. Bei den ersten Loebner Wettbewerben von 1991 bis 1993 gewann Joseph Weintraub mit seinem Programm „PC Therapist“ die Bronzemedaille. Aus dem Namen geht bereits die deutliche Anlehnung an ELIZA hervor. Angeblich aus Enttäuschung über die Leistungsfähigkeit ELIZAs begann der Psychologe Weintraub 1986 mit der Entwicklung seines Programms. Abgesehen von der Möglichkeit, Antworten in gesprochener Sprache zu synthetisieren scheint der wesentliche Unterschied zu ELIZA weniger auf der verwendeten Programmiertechnik zu beruhen, als auf Vokabular und „Persönlichkeit“ des Chatbots.²² Da Weintraub sein Programm auch heute noch vermarktet, sind die Einzelheiten über die Funktionsweise nicht zugänglich. Die Tatsache, dass „PC Therapist“ auch 1995 die Bronzemedaille gewann, zeigt, wie vergleichsweise langsam die Entwicklung moderner Chatbot-Systeme vonstatten ging.

1994 ging Thomas Whalen mit TIPS als Sieger aus dem Loebner Wettbewerb hervor. TIPS orientierte sich stark an der Architektur klassischer Datenbanksysteme, ein Aufbau, der den Anforderungen eines thematisch eingeschränkten TTs entgegenkam. Als im folgenden Jahr der uneingeschränkte TT eingeführt wurde, erwies sich die Systemarchitektur von TIPS als zu unflexibel.

Die Bronzemedaille von 1996 gewann Jason Hutchens mit HeX. HeX analysiert die Nutzereingabe Satz für Satz. In einem ersten Schritt wird nach Schlüsselwörtern gesucht. Findet sich für eines dieser Schlüsselwörter eine ausformulierte, im laufenden Gespräch noch nicht gebrauchte Antwort in einer Teildatenbank, so wird diese ausgegeben. Ist dies nicht der Fall, wird der Satz

²¹ ebd.

²² vgl. Hutchens (1997:9), Saygin, Cicekli, Akman (2000:37) „PC Therapist“ gilt als „Meister des non sequitur“.

mit einer Reihe von möglichen „Tricks“ wie mathematischen Fragen oder Nonsense-Eingaben abgeglichen. Hutchens hatte die Protokolle der vorangegangenen Wettbewerbe gesichtet und daraus mögliche Fallen abgeleitet. Führen diese beiden Schritte nicht zum Erfolg, wird ein weiteres Modul, MegaHal, aktiviert. MegaHal, der eigentliche Chatbot im System, kann Sätze analysieren und eigenständig Antworten generieren. Dazu modelliert Hutchens Sprache in Form einer Markov Kette. Mit Hilfe dieser Kette lassen sich Zeichenketten generieren. So entsteht zunächst nur ein „babble generator“²³ Um die generierten Antworten in Zusammenhang mit der Eingabe zu bringen, wird diese auf Schlüsselwörter untersucht. Verschiedene Algorithmen sorgen dann für die Generierung einer Antwort auf der Basis dieser Schlüsselwörter. Im Unterschied zu den vorangegangenen Siegern ist MegaHal / HeX selbständig lernfähig.²⁴

Auch der Sieger von 1998 und 1999²⁵, FRED bzw. Albert One²⁶, „lernt“. Robby Garner arbeitete 15 Jahre an diesem System, bevor er es zum Loebner Wettbewerb anmeldete. FRED kann als natürlichsprachige Schnittstelle zu unterschiedlichsten Programmen eingesetzt werden und unterscheidet sich daher von anderen Loebner Teilnehmern. FREDs Wissen beruht auf einer großen Datenbank mit Sätzen, die ihm bisher im Gespräch begegnet sind. Diesen Sätzen sind Antworten zugeordnet. Besteht die Eingabe aus einem unbekanntem Satz, wird in der Datenbank ein möglichst ähnlicher Satz gesucht und die entsprechende Antwort ausgegeben. Später können diesen neuen Sätzen Antworten zugeordnet werden. Dies geschieht allerdings durch den Programmierer. FRED generiert keine eigenen Antworten.

In den Jahren 2000 und 2001 gewann Richard Wallace mit ALICE die Bronzemedaille. Wir werden im Folgenden detailliert auf dieses Programm eingehen.

²³ Hutchens (1997:14)

²⁴ vgl. Hutchens (1997:19f)

²⁵ 1997 erhielt David Levy die Bronzemedaille. Über sein System „Converse“ liegen jedoch kaum Informationen vor.

²⁶ Dieses System wird in allen Quellen als FRED bezeichnet. Lediglich auf der Homepage seines Entwicklers heißt es „Albert One“. (www.cybermedia.com/new%20folder/Robby%20Garner/)

5 Intelligente Sprachverarbeitung?

Sprachfähigkeit ist ein wesentlicher Bestandteil menschlicher Intelligenz. Auch wenn Dissoziationen zwischen Sprachbeherrschung und allgemeiner Intelligenz zu beobachten sind, entwickeln sich Sprache und Intelligenz doch in enger Wechselwirkung.²⁷ Diesen Sachverhalt legte bereits Turing seinem Imitationsspiel zu Grunde. Auch für die KI-Forschung stellt sich in diesem Zusammenhang die Frage, ob Maschinen wie Menschen denken oder zumindest wie Menschen reden können und wie diese Fähigkeit auf Maschinen übertragen werden kann. Um Computer zum Sprechen zu bringen, muss - so die gängige Meinung der KI-Forschung - zunächst geklärt werden, wie Sprache strukturiert ist und wie sie funktioniert. Es erscheint notwendig, natürliche Sprache auf theoretische und mathematische Modelle zurückzuführen, die dann dem Computer zugänglich gemacht werden können.

Nach Weizenbaums ELIZA entstanden (im Bereich der „ernsthaften“ KI) erste Programme, mit denen zunehmend anspruchsvollere Sprachverarbeitungsaufgaben ausgeführt werden konnten.²⁸ Ein frühes System war Shrdlu, das in einer fiktiven Welt Bauklötze auf Anweisung manipulieren und Fragen zum Zustand dieser Welt beantworten konnte.²⁹ Anwendungen wie OCR (optical character recognition) zur Erkennung von Schriftzeichen, natürlichsprachige relationale Datenbanksysteme, Diktiersysteme oder Sprachsteuerungen für Computer und Mobiltelefone. Diese Systeme sind jedoch auf einzelne Bereiche der Sprachverwendung beschränkt. Bislang existiert keine Lösung, die den Stand menschlicher Sprachbeherrschung erreicht. Die Forschung, so Menzel, brachte auf diesem Gebiet vor allem „Einsicht in die ungeheure Komplexität und Vielschichtigkeit des Phänomens Sprachbeherrschung“³⁰

Vor diesem Hintergrund treten seit Anfang der 90er Jahre verstärkt die Chatbots auf, die von ihren Entwicklern gern mit Attributen wie Intelligenz oder

²⁷ vgl. Menzel 2000:643

²⁸ vgl. Menzel 2000:644

²⁹ Einen Überblick über die frühen „sprachverstehenden“ Systeme geben Hutchens, Weizenbaum (1976) und Raphael (1976)

³⁰ Menzel 2000:644

Sprachverstehen ausgezeichnet werden.³¹ So wirbt z. B. die Firma Kiwilogic mit den Worten „Ein Lingubot³² kann Fragen in natürlicher Sprache verstehen und direkt beantworten“ für ihre Chatbotsoftware. Mit derartigen Aussagen machen die Entwickler somit Versprechungen auf einem Gebiet, auf dem die etablierte Forschung nach zahlreichen Enttäuschungen Zurückhaltung übt.³³

Ist den „Spielkindern“ der KI mit der Entwicklung von Chatbots ein Quantensprung in der Mensch-Maschine-Kommunikation gelungen oder täuschen diese Programme auf geschickte Weise (sprachliche) Intelligenz vor? Wir gehen dieser Frage nach, indem wir zunächst die Charakteristika natürlicher Sprache und die daraus resultierenden Schwierigkeiten ihrer maschinellen Verarbeitung zusammenfassen. Anschließend wird am Beispiel von ALICE untersucht, wie Chatbot-Programmierer versuchen, diese Schwierigkeiten zu meistern.

5.1 Vom Verstehen natürlicher Sprache

Unter natürlicher Sprache werden im Allgemeinen historisch entwickelte, regional und sozial geschichtete Sprachen verstanden. Natürliche Sprache unterliegt somit Sprachwandelprozessen. Sie ist durch verschiedene Formen von Ambiguität gekennzeichnet. Im Gegensatz dazu zeichnen sich künstliche Sprachsysteme, wie beispielsweise Programmiersprachen durch ihre Unveränderlichkeit und vor allem ihre kompromisslose Eindeutigkeit aus.³⁴ Im Vergleich mit solchen künstlichen Sprachen treten die Charakteristika natürlicher Sprache deutlich hervor:

- Natürliche Sprache bietet auf der Basis einer begrenzten Menge von Elementen unendlich viele Kombinationsmöglichkeiten. Morpheme werden zu neuen Wörtern kombiniert, die auch verstanden werden, wenn sie noch nicht lexikalisiert

³¹ vgl. dazu z. B. die Dokumentation zu ALICE: www.alicebot.net

³² Lingubot ist eingetragenes Warenzeichen der Firma Kiwilogic

³³ vgl. Lessmöllmann 2000

³⁴ Programmiersprachen verfügen über einen begrenzten „Wortschatz“ (Befehle, Kommandos), der im Gegensatz zu den Wortbildungsprozessen natürlicher Sprachen nicht durch Kombination vorhandener Elemente erweitert werden kann. Jeder Befehl hat genau eine Bedeutung und erfordert eine bestimmte Syntax. Bereits minimale Abweichungen führen zu Fehlermeldungen: Der Computer „versteh“ nicht mehr.

sind; die Möglichkeiten der Satzbildung sind trotz syntaktischer Beschränkungen immens.

- Natürliche Sprache unterscheidet eine Vielzahl von Sprechhandlungen.³⁵ Daraus resultiert pragmatische Ambiguität: Ist der Satz „Ist das Fenster offen?“ also eine Frage, eine indirekte Aufforderung, ein Vorwurf [...]?
- Natürlichsprachige Äußerungen folgen sprachökonomischen Prinzipien und sind daher häufig unterspezifiziert, elliptisch. Der Prozess des Verstehens beinhaltet somit die Rekonstruktion des impliziten Informationsgehalts.³⁶
- Natürlichsprachige Ausdrücke können im uneigentlichen Sinn gebraucht werden: Metapher und Metonymie, Ironie

Neben diesen Schwierigkeiten besteht das Hauptproblem der maschinellen Verarbeitung in der Ambiguität natürlicher Sprache:³⁷

- Lexikalische Ambiguität (Polysemie und Homonymie)
- Strukturelle / syntaktische Ambiguität: Wie ist der Satz „Time flies like a banana“ zu analysieren?
- Referentielle Ambiguität: Pronomen lassen sich nicht allein aufgrund der sprachlichen Umgebung zuordnen: Sie nahm den Kuchen vom Teller und aß ihn.
- Pragmatische Ambiguität (s.o.)

Das menschliche Sprachverstehen erweist sich angesichts dieser Merkmale natürlicher Sprache als ausgesprochen fehlertolerant und leistungsfähig. Für den Menschen sind sprachliche Äußerungen keine isolierten Gebilde. Sie stehen in einem sprachlichen und situativen Kontext, haben Bezug zu einer außersprachlichen Realität. Diese Einbettung ist Voraussetzung für die Mühelosigkeit und das Tempo der Sprachverarbeitung. Görz fasst dies wie folgt zusammen:

„Damit ist die intendierte Bedeutung jeder Äußerung von der Folge der ihr vorausgegangenen Ereignisse und Äußerungen abhängig, von der

³⁵ vgl. Görz 1988:7

³⁶ vgl. Menzel 2000:646

³⁷ vgl. Görz 1988:8. Bei der Verarbeitung gesprochener Sprache vervielfachen sich die Probleme durch individuelle Unterschiede in der Aussprache, Verschleifung der Wortgrenzen etc.

Kommunikationssituation (Ort, Zeit, beteiligte Personen, Umgebung, Anlaß und Zweck der Äußerung) und von der Gesamtheit des beim Hörer vorausgesetzten Vorwissens.³⁸

Der Mensch scheint über eine innere Repräsentation der Welt zu verfügen, die es ihm ermöglicht, beim Verstehensprozess Bedeutung zu rekonstruieren, Lücken zu füllen oder im Falle mehrerer Varianten die plausibelste zu wählen. Dieses Wissen um Dinge und Sachverhalte, diese Idee von der Beschaffenheit der Welt, ist für das Verstehen natürlicher Sprache unerlässlich. Das bedeutet keineswegs, dass Wissen um die Struktur von Sprache bedeutungslos wäre. Denn, und damit greifen Strube et al ein Argument Chomkys auf: „Wäre es nun tatsächlich so, daß wir Sprache nahezu ausschließlich aufgrund unserer Kenntnis der Welt verstehen, könnten wir nichts ausdrücken, was wir nicht schon wissen.“³⁹ Es ist das Zusammenspiel von sprachlichem „Fachwissen“ und Wissen über die Welt, die Menschen das Verstehen ermöglicht.

Während es gelungen ist, den strukturellen Aspekt von Sprache, z.B. beim Parsing syntaktischer Strukturen, maschinell zu verarbeiten, bereitet die Wissensrepräsentation noch immer große Probleme. Es stellt sich die Frage, wie man das eher diffuse Weltwissen von Menschen für den Computer modellieren kann, zumal nicht geklärt ist, wie die Repräsentation beim Menschen aussieht. Bisher wurde oft auf die Repräsentation in semantischen Netzwerken und/oder Skripts, für die Darstellung stereotyper Ereignisabläufe zurückgegriffen. Neben diesen Ansätzen der symbolischen KI, versucht die Forschungsrichtung des Konnektionismus mittels neuronaler Netzwerke die Prozesse im menschlichen Gehirn abzubilden. Der Aufwand dabei ist immens. Es zeigen sich dennoch Teilerfolge.⁴⁰ Die realisierten Systeme beschränken sich aber auf thematisch stark eingegrenzte Gebiete. Die Lösungsvorschläge sowohl der symbolischen KI als auch des Konnektionismus, so Steup, „unterliegen dem Problem der kombinatorischen Explosion beim Übergang von kleinen Teilproblemen hin zur

³⁸ Görz 1988:8

³⁹ Strube et al 2000:51 führt Expertensysteme, maschinelle Übersetzung etc. an.

⁴⁰ Görz 2000 bietet einen ausführlichen Einblick in unterschiedliche Lösungsansätze und stellt einzelne Systeme vor.

realen Welt“.⁴¹ Von der Entwicklung eines Systems mit allgemeiner, menschenähnlicher Sprachfähigkeit ist die „klassische“ KI auch heute noch weit entfernt.

5.2 Die Funktionsweise von Chatbots - ALICE

Die heutigen Chatbots stehen auch technisch in der Tradition von ELIZA. Sie sind in der Regel genau wie ihr Vorbild modular aufgebaute Reiz-Reaktions-Systeme, die den sprachlichen Input mit einer internen Musterdatenbank abgleichen und entsprechende Antworten ausgeben.

Das Herzstück des Chatbots ist auch heute die so genannte Wissensdatenbank, der Teil des Systems, in dem Erkennungsmuster, Stichwörter, Antworten gespeichert sind. Das eigentliche Programm steuert den Gesprächsablauf, koordiniert also Eingabe und Ausgabe, die Aktivierung der Wissensdatenbank und ggf. weitere Module wie die Ausgabe gesprochener Sprache. Vor allem kommerzielle Systeme stellen außerdem einen Editor zu Verfügung, der dem Anwender Aufbau und Pflege der Wissensdatenbank erleichtert. Alle Chatbots verfügen über eine Protokollfunktion, die sämtliche Dialoge speichert und dem Entwickler zur Auswertung bereitstellt.

Die einzelnen Systeme unterscheiden sich allerdings in Flexibilität und Größe ihrer „Wissensdatenbanken“ sowie in der der Leistungsfähigkeit ihrer Steuerungsprogramme: Während einige Chatbots nur schriftsprachlich kommunizieren, kann zum Beispiel ALICE gesprochene Antworten generieren und „lernt“ zur Zeit auch Eingaben in gesprochener Sprache zu „verstehen“. Es gibt Systeme, die, wie bereits ELIZA, Antworten aus Rümpfen generieren. Andere hingegen wählen aus einer Reihe von ausformulierten Antworten per Zufallsprinzip aus. Einige Chatbots vergleichen nur den Teil einer Eingabe mit der Musterdatenbank, der vor einem Satzzeichen steht und lassen eventuell folgende Sätze bei der Antwort unberücksichtigt. Andere analysieren alle Teile der Eingabe und kombinieren die Antworten auf den jeweiligen Satz miteinander. Chatbots unterscheiden sich außerdem in ihrem „Gedächtnis“, also darin, ob Teile von Eingaben in Variablen wie Alter, Name, Beruf,

⁴¹ Steup 1996:114

Gesprächsthema gespeichert und im Verlauf des Chats wieder aufgerufen werden können. Nicht zuletzt stellt sich die Frage der Lernfähigkeit. Die meisten Chatbots führen Protokoll über ihre Gespräche und schlagen darauf beruhende neue Erkennungsmuster oder Stichwörter vor. Ob diese dann in die Wissensdatenbank aufgenommen werden, entscheidet ein Mensch. Dieses Verfahren wird *supervised learning* (überwachtes Lernen) genannt. Es ist jedoch fraglich, ob hierbei nicht eher der programmierende Mensch aus seinen Fehlern lernt als die Maschine. Es gibt jedoch auch Systeme, die anhand der Protokolle selbständig ihre Datenbank erweitern können wie z.B. Fred und MegaHAL. Einige Bots können externe Datenquellen nutzen, wenn für die Eingabe kein Erkennungsmuster vorliegt.

Zusammenfassend lässt sich hinsichtlich dieser Unterschiede festhalten, dass die am Loebner-Wettbewerb teilnehmenden Systeme eher flexibler programmiert sind, während die kommerziellen Lösungen auf strikte Frage-Antwort-Schemata aufzubauen scheinen. Dies mag auch daran liegen, dass beim Loebner-Preis seit einigen Jahren das Themenfeld nicht mehr eingeschränkt wird, während kommerzielle Anbieter sich durch die Festlegung auf bestimmte Produkte und Services ohnehin thematisch beschränken und zudem „unpassende“ Antworten ihres Chatbots von vornherein ausschließen wollen.

Zur Zeit gilt ALICE mit zwei Bronzemedailles im Loebner Contest als erfolgreichster Chatbot. Da der Entwickler Dr. Richard Wallace Programm und Wissensdatenbank von Anfang an als Open Source veröffentlicht hat, existiert heute eine ganze Reihe von ALICE-Klonen⁴² im Internet. Zudem arbeiten mehrere Entwickler parallel an diesem Projekt, was zu schnellen Fortschritten in der Entwicklung führt.⁴³

ALICE steht für Artificial Linguistic Computer Entity. Auch ALICE ist modular aufgebaut und besteht aus einer ganzen Reihe von Programmen, die in einer Client-Server-Architektur miteinander verbunden sind. Uns interessiert dabei der Aufbau der Wissensdatenbank.⁴⁴ ALICE's Kern ist in AIML

⁴² Eine Liste von ALICE-Klonen findet sich im Literaturverzeichnis

⁴³ Wir beziehen uns im Folgenden auf die englischsprachige ALICE-Version, da hierzu eine ausführliche Dokumentation im Internet vorliegt.

⁴⁴ Eine ausführliche Beschreibung der ALICE Architektur und Funktionsweise findet sich unter <http://birch.eecs.lehigh.edu/alice/readme.html> und unter <http://www.german.alicbot.com>

(Artificial Intelligence Markup Language), einer XML-Sprache, die in ihrer Struktur HTML ähnelt, geschrieben. Dies hat zwei Vorteile: Zum einen ist AIML intuitiv und leicht erlernbar, zum anderen ist sie um neue Befehle (Tags) erweiterbar; eine Eigenschaft, die klassischen Programmiersprachen zunächst fehlt.

Die Wissensdatenbank besteht aus einer langen Liste so genannter *categories*. Eine *category* wiederum setzt sich zusammen aus einem *pattern*, also einem Erkennungsmuster, und einem *template*, der Antwort:

```
<category>
<pattern> HALLO </pattern>
<template> Hallo, ich bin ALICE. Wie geht es dir? </template>
</category>
```

Allein diese vier Zeilen AIML bewirken, dass der Chatbot auf die Eingabe „Hallo“ mit der Antwort „Hallo, ich bin ALICE. Wie geht es dir?“ reagiert.

AIML erlaubt einen Platzhalter (Wildcard) pro Erkennungsmuster. Mit diesem Zeichen lassen sich die Muster verallgemeinern, der Chatbot reagiert nur auch auf Eingaben wie „Hallo Alice“ oder „Hallo Du“.

```
<category>
<pattern> HALLO * </pattern>
[...]
```

Auf andere Begrüßungen wie „Guten Tag“ erfolgt dagegen keine Reaktion. Hierfür müssen jeweils eigene *categories* geschrieben werden. Damit wird bereits deutlich, dass das Erstellen einer Wissensdatenbank mit einem hohen Maß an detaillierter, antizipierender Schreibearbeit verbunden ist.

Mit den Tags `set_X` und `get_X` erhält ALICE ein Gedächtnis. Der Befehl `set` speichert einen bestimmten Teil der Eingabe in einer Variablen, `get` ruft den Inhalt dieser Variable erneut auf. Auf diese Weise lassen sich recht einfach auf den ersten Blick verblüffende Effekte erzielen. Durch die folgenden *categories*

lässt sich, vor allem in Verbindung mit weiteren Standarderkennungen bereits eine einfache Form von Bewusstsein des „Gesagten“ vortäuschen.

```
<category>
<pattern> MEINE SCHWESTER * </pattern>
<template><set_sie>deine Schwester</set_sie> </template>
</category>
<category>
<pattern>WOFÜR STEHT SIE </pattern>
<template> „Sie“ steht für „<get_sie/>“ </template>
</category>
```

Mensch: Meine Schwester heißt Margot. Sie hat gestern einen Unfall gehabt.

Alice: Oh, das tut mir leid. Hat sie sich verletzt?

Mensch: Nein. Hat sie nicht. Wofür steht sie?

Alice: „Sie“ steht für deine Schwester.

Dies ist für eine ganzen Reihe von Variablen wie Benutzername, Wohnort, Beruf, Thema oder Sprache vorgesehen. Erweiterungen um neue Tags sind jederzeit und durch jeden möglich. Es besteht ferner die Möglichkeit, ALICE aus einer Reihe von Antworten zufällig auswählen zu lassen. Dies empfiehlt sich vor allem bei Erkennungsmustern, die in einem Dialog häufiger auftreten könnten, damit der Chatbot sich nicht wiederholt:

```
<category>
<pattern>BIST DU * </pattern>
<template>Ich weiss nicht, ob ich <star/> bin.
Ich bin <random>
<li>ein Computer</li>
<li>ein von Dr. Wallace entwickelter Chatterbot</li>
<li>ein gigantisches elektronisches Gehirn</li>
<li>klüger als du</li>
```

```
</random> </template>
</category>
```

Das Tag `<star/>` wiederholt die letzte Zeichenkette, also das letzte Wort der Eingabe. Es gibt mehrere Tags, die sich zur Manipulation von Zeichenketten einsetzen lassen. So ist es z.B. möglich mit `<recallu>` die letzte Eingabe des Nutzers zu wiederholen. Weitere Funktionen sind der Ersatz von Personalpronomen, Zufallsantworten auf Ja/Nein-Fragen oder Erzählen von Geschichten auf der Basis einer gesonderten Geschichten-Datei.

Die Wissensdatenbank von ALICE stellt sich als eine große Textdatei mit AIML-Tags dar. Die Download-Version umfasst inzwischen ca. 25.000 Kategorien. Damit ist die Wahrscheinlichkeit, dass ALICE auf eine Nutzereingabe angemessen reagieren kann, für einen Chatbot schon relativ hoch. In Verbindung mit dem Wechsel der Personalpronomen, dem Einsatz von Pronomen (`<set_sie>`) und den Gedächtnisfunktionen, lassen sich recht ausgefeilte Dialoge führen. Und dies ist nach Wallace das Hauptziel von ALICE: „ALICE has always been aimed toward keeping the client talking as long as possible, not necessarily giving any correct information along the way.“⁴⁵

Aber kann ALICE denken? Wallace antwortet auf diese Frage mit „more or less“⁴⁶. Tatsächlich gibt es ein `<think>` Tag:

```
<category>
<pattern> MY NAME IS * </pattern>
<template>
<think> <set_name> <star/> </set_name> </think>
<get_name/>, nice to meet you!
</template>
</category>
```

`<think>` trennt lediglich den Vorgang des Wertsetzens für eine Variable von der Ausgabe. Inwiefern sich das von einer einfachen set / get Kombination

⁴⁵ alicebot.org/articles/bush/wallaceConversation.html

unterscheidet, geht aus der Dokumentation nicht hervor. Mit einem wie auch immer gearteten Denkprozess scheint <think> jedoch wenig gemein zu haben. In der deutschen Dokumentation zu ALICE findet sich in Bezug auf das Denkvermögen ein anderer Hinweis. Da ALICE auch durch Satzzeichen unterteilte Eingaben zwar vollständig aber in Teilen analysiert, kann es bei der Antwort zur Kombination von zwei oder mehr *templates* kommen:

- Ich wette, du bist schwul
- Eigentlich bin ich kein Spielertyp. Eigentlich habe ich als Maschine kein Bedürfnis nach Sex.

Der erste Satzteil enthält das Stichwort wetten und erzeugt einen entsprechenden Antwortsatz. Der zweite Teil der Eingabe passt auf die Stichwörterkennung schwul und gibt ebenfalls eine entsprechende Antwort aus. Was von den ALICE Programmierern als Anzeichen von Denken gedeutet wird, ist unseres Erachtens die strikte Befolgung des Suchalgorithmus, der dem Chatbot zugrunde liegt. Dabei entstehen zufällig Kombinationen von Antworten, die einen gemeinsamen Sinn ergeben. Diese sind zwar vom Programmierer nicht geplant bzw. antizipiert worden, ein Zeichen für Denkfähigkeit ist darin dennoch nicht zu sehen.

Stellt man nun ALICE den vorab geschilderten Anforderungen bei der maschinellen Verarbeitung natürlicher Sprache und den Anforderungen der KI-Forschung an sich selbst gegenüber, zeigt sich deutlich, dass ALICE sich wie alle anderen Chatbots auch außerhalb dieses Forschungsparadigmas befindet. ALICE folgt in unterschiedlicher Hinsicht einem grundlegenden Prinzip: Minimalismus.

Zum einen sollen Aufwand und Schwierigkeit bei der Programmierung möglichst gering bleiben. Daher wurde die vergleichsweise einfache Sprache AIML gewählt. Zum anderen ruht auch die Gestaltung der Wissensbasis auf einem minimalen theoretischen Gerüst. Um einen erfolgreichen Chatbot zu entwickeln, benötigt man weder ausgefeilte Sprachtheorien, noch komplexe

⁴⁶ Wallace: www.alicebot.net/thinking.html

Wissensrepräsentationen oder Lerntheorien. Dieses Credo der ALICE-Gemeinde wird in einigen Zitaten Wallace's deutlich:

„You don't need a complex theory of learning, neural nets, or cognitive models to explain how to chat within the limits of Alices's 25,000 categories. Our stimulus-response model is as good a theory as any other for these cases, and certainly the simplest. [...] If there is any room left for ‚higher‘ natural language theories, it lies outside the map of the Alice brain. [...] The work of those seeking to explain natural language in terms of something more complex than stimulus response will take place beyond our frontier, increasingly in the hinterlands occupied by only the rarest forms of language. [...] Our territory of language already contains the highest population of sentences that people use.“⁴⁷

Auch die „kombinatorische Explosion“, die der gängigen KI beim Übergang von ihren Modellen zur realen Welt Schwierigkeiten bereitet, stellt für ALICE zunächst kein Problem dar. Wallace verweist in diesem Zusammenhang auf das Zipfsche Gesetz⁴⁸, nach dem die Anzahl der theoretischen Kombinationsmöglichkeiten weitaus höher als die Anzahl der tatsächlich genutzten Kombinationen ist: „Considering the vast size of the set of things people could possibly say, that are grammatically correct or semantically meaningful, the number of things people actually do say is surprisingly small.“⁴⁹ Folgt man dieser Argumentation, erscheint es fast möglich, die tatsächlichen gebrauchten Äußerungen annähernd komplett zu antizipieren.

Wer sich mit Sprache und Sprachverarbeitung beschäftigt, kann in ALICE kaum einen echten Durchbruch auf dem Gebiet der maschinellen Sprachverarbeitung bzw. des maschinellen Sprachverstehens sehen. Es lässt sich jedoch nicht abstreiten, dass sich die Dialoge mit ALICE als erstaunlich flüssig erweisen, das System durchaus erfolgreich ist. Und das, obwohl sich seit ELIZA

⁴⁷ Wallace: www.alicebot.net/drwallace.php?column=9

⁴⁸ 1935 erkannte G.K. Zipf den gesetzmäßigen Zusammenhang zwischen der Verwendungshäufigkeit von Wörtern in einzelnen Texten bzw. bei einzelnen Sprechern / Autoren und ihrem Rangplatz auf einer Liste ihrer generellen Auftretungshäufigkeit. Die Formel bzw. die Zipfsche Verteilung sind universell. Es besteht ferner eine Korrelation zwischen der Länge eines Wortes und seiner Häufigkeit. Dies hängt laut Zipf mit dem ökonomischen Prinzip des geringsten Kraftaufwands zusammen.

⁴⁹ Wallace: www.alicebot.net/drwallace.php?column=4

nicht viel am Verfahren geändert hat. Und so kann Wallace sein einfaches Reiz-Reaktions-System als Erfolg verkaufen und sich einen Seitenhieb auf die traditionelle KI erlauben: „The academic world-view of AI in particular rejects any ‚ELIZA-like‘ approach to natural language as too simplistic. Unfortunately for them it happens to be the one theory that works.“⁵⁰

5.3 Tricks

Zumindest bis zu einem gewissen Punkt scheint Wallace’s minimalistischer Ansatz in Kombination mit einem Reiz-Reaktions-System erfolgreich zu sein. Beim Loebner Contest wurde bisher jedes Programm als solches identifiziert. Doch auch für den Gewinn einer Bronzemedaille ist es nicht ausreichend, eine möglichst große Wissensdatenbank mit einigen technischen Feinheiten, wie der Umstellung von Personalpronomen zu erzeugen. Die Entwickler arbeiten mit einer ganzen Reihe von Tricks, die ihren Chatbot möglichst menschlich erscheinen lassen sollen.⁵¹

Antwortzeit

Die meisten Chatbots arbeiten aufgrund leistungsstarker Computer sehr schnell. Da sie jedoch einen Menschen simulieren, der seine Eingaben auf einer Tastatur tippen muss, erfolgt die Ausgabe verlangsamt. Andererseits kann eine zu lange Verarbeitungszeit durch kompliziertes Parsen oder immense Datenmengen ebenso problematisch sein. Hutchens hat sein System daher mit einer Time-out-Routine ausgestattet, die nach einer bestimmten Zeit eine belanglose Auffangantwort ausgibt.⁵²

Fehler

Wird unter Zeitdruck getippt, kommt es zu Fehlern. In einigen Chatbots sind entsprechend Tippfehler eingebaut. Chatbots machen auch Fehler bei arithmetischen Berechnungen. Sie können aber auch Fehler in der Eingabe

⁵⁰ Wallace: alicebot.org/articles/bush/wallaceConversation.html

⁵¹ Eine ausführliche Zusammenstellung von Tipps und Tricks zur Erstellung von Chatbots findet sich unter www.abenteuermedien.de/jabberwock/faq_de.html

⁵² vgl. Hutchens (1997:17)

erkennen und ggf. kommentieren. ALICE beispielsweise bemerkte während eines Chats, wir hätten „he’s“ statt „his“ verwendet und fragte, ob sie das richtig erkannt habe.

Gesprächsthemen

Erfolgreiche Chatbots sind in der Lage, Gesprächsthemen einzubringen oder kleine Geschichten zu erzählen, um so den Chat in Bereiche zu lenken, in denen sie über Erkennungsmuster verfügen. Denn schließlich ist es unmöglich, das Programm auf jede erdenkliche Nutzereingabe vorzubereiten. An diesem Punkt greift Wallace’s Auslegung des Zipfschen Gesetz. In der Tat scheinen Nutzer ihre Möglichkeiten nicht auszuschöpfen. Whalen setzte bei seinem System von 1995 (Joe) auf die von Dale Carnegie ermittelten Gesprächsthemen, die Fremde bei ihrer ersten Begegnung in der Regel anschneiden.⁵³ Whalen scheiterte unter anderem daran, dass die Juroren des Loebner Wettbewerbs sich den Probanden nicht wie Fremde näherten, sondern ohne Rücksicht auf kulturelle Konversationsmuster Fragen stellten. Außerhalb der spezifischen Testsituation mag diese Strategie durchaus erfolgreich sein. Entwickler wie Hutchens und Wallace erweitern das „Wissen“ ihrer Chatbots um enzyklopädisches Wissen, das aus dem Internet gewonnen wird. Sie versorgen ihre Programme mit bekannten Zitaten und fremdsprachlichen Floskeln. Viele Chatbots, vor allem kommerziell eingesetzte Programme, verfügen über ein Spezialgebiet, in dem sie über sehr viele Erkennungsmuster und ausgefeilte Antworten verfügen. Dies allein ist jedoch nicht ausreichend. Um menschliches Gesprächsverhalten zu simulieren, bedarf es zusätzlich einer großen Zahl von Mustern für allgemeine – auch phatische – Kommunikation.

non sequitur

Auch dem Programm mit der größten Wissensdatenbank wird es nicht gelingen, für jede Eingabe ein passendes Erkennungsmuster zu finden. Entscheidend für die Leistungsfähigkeit eines Chatbots ist vor allem, wie er mit solchen Eingaben umgeht. Die vergangenen Loebner Wettbewerbe haben

⁵³ Nach Whalen sind dies, in entsprechender Reihenfolge: Der eigene Name, der Wohnort, ehemalige Wohnorte, gemeinsame Bekannte, Wetter, Sport, Politik, Bücher – Musik – Fernsehen und Kino, Hobbies.

gezeigt, dass Auffangantworten wie „Ich weiß nicht.“ oder „Was meinen Sie mit X.“ von den Juroren kaum toleriert werden.⁵⁴ Im Gegensatz dazu bereiten wirre, enigmatische, aus dem Zusammenhang fallende Repliken keine Probleme. Die Juroren zeigten sich gewillt, auf diese einzugehen, die Konversation lief weiter.

Wiederholungen

Antwortet ein Programm auf dieselbe Eingabe immer mit einer einzigen Standardantwort, wird es schnell entlarvt. Es ist unwahrscheinlich, dass ein Mensch auf eine zum x-ten Mal in Folge gestellte Frage („Was ist der Sinn des Lebens?“) immer geduldig mit derselben Antwort („42“) reagiert. Die Entwickler streben daher an, für ein Erkennungsmuster eine ganze Reihe von Antworten zur Verfügung zu haben, aus denen das System zufällig auswählt. Ein guter Chatbot wird zudem versuchen, das Gespräch in eine andere Richtung zu lenken („42, ist es nicht schade, dass Douglas Adams tot ist?“).

Persönlichkeit

Nicht zuletzt scheint es erfolgversprechend, dem Chatbot eine Persönlichkeit, eine virtuelle Identität zu verschaffen. Dies muss keineswegs bedeuten, dass es sich dabei um eine menschliche Figur handelt. Im Gespräch werden die Programme noch immer recht schnell entlarvt, so dass sich das Vortäuschen eines Menschen eher negativ auf die Akzeptanz auswirkt. Allerdings sollte man sich bewusst sein, dass die Art, in der Menschen miteinander kommunizieren, ihre Wortwahl, die Codierung von Stimmungen usw. viel über die Persönlichkeit aussagen. Es wäre also von Vorteil, den Chatbot in dieser Hinsicht konsistent zu gestalten. Die Tatsache, dass der durchschnittliche Chatbot im Regelfall mit mehr unbekanntem als vorgesehenen Eingaben konfrontiert wird, macht dies zu einem schwierigen Punkt. In den Tipps zur Chatbotgestaltung⁵⁵ heißt es dazu:

„Machen Sie Ihren Chat-Roboter absonderlich und verrückt! [...] Je unsinniger Ihr Bot reagiert, desto mehr der Anwender sind davon überzeugt, mit einem

⁵⁴ Whalen scheiterte 1995 auch, weil sein Programm Joe zu oft vorgab, „keine Ahnung zu haben“. Joe war als ungebildeter Hausmeister konzipiert, der nie Zeitung las, keinen Fernseher besaß und nur nachts arbeitete. (vgl. Whalen)

⁵⁵ s. www.abenteuermedien.de/jabberwock/faq_de.html

Menschen zu sprechen. [...] Der Bot wird als psychotisches Individuum wahrgenommen – und dies ist ein menschliches Attribut.“

Diesen Trick verwendete schon PARRY. Und hinsichtlich der Akzeptanz von non sequitur Repliken scheint er erfolgsversprechend. Ob sich aber die Inhaber von e-commerce Seiten einen neurotischen Chatbot als Gastgeber vorstellen können bleibt fraglich.

6 Fazit

„I believe that in about fifty years' time, it will be possible to programme computers with a storage capacity of about 10^9 , to make them play the imitation game so well, that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning.“⁵⁶

Turings Prognose hat sich bis heute nicht bewahrheitet. Noch immer sind die entsprechenden Computersysteme weit davon entfernt, fünf Minuten lang menschlich zu wirken. Hutchens Kommentar zur Leistungsfähigkeit der Chatbots ist eindeutig: „One cannot help but think that he would have been disappointed, and that his comments would have had a great impact on the future directions of such works“⁵⁷. Turing starb 1954, rund zehn Jahre bevor die ersten Programme wie ELIZA entstanden.

Die Beobachtungen in der Welt der Chatbots haben gezeigt, dass sich seit ELIZA das Grundprinzip der Programme nicht wesentlich verändert hat. Noch immer wird die (getippte) Nutzereingabe mit einer Wissensdatenbank abgeglichen, werden überwiegend vorformulierte Antworten auf dem Bildschirm ausgegeben. Abgesehen von selbstlernenden Bots wie MegaHal, die auf einer tatsächlichen Modellierung von Sprache beruhen, gilt weiterhin, dass die Sprachfähigkeiten eines Systems letztlich den Fähigkeiten und Fertigkeiten des Programmierers entspricht und seinem Geschick, mögliche Eingaben zu antizipieren. Allerdings haben sich seit ELIZA mit zunehmend leistungsfähigeren

⁵⁶ Turing (1950:442)

⁵⁷ Hutchens (1997: 2)

Computern die Rahmenbedingungen verbessert. Dank großer Mengen verfügbaren Speicherplatzes und schnellen Prozessoren lassen sich heute mit einer ELIZA ähnelnden Architektur wesentlich bessere Ergebnisse erzielen als in den 60er Jahren. Die Entwickler profitieren also nicht zuletzt von den Weiterentwicklungen im Hardwarebereich. Der Erfolg eines Bots wird weniger von der technischen Machbarkeit als von der damit verbundenen Fleißarbeit, möglichst viele Eingaben zu antizipieren begrenzt.

Doch auch dem leistungsfähigsten Chatbot gelingt es unserer Ansicht nach nicht, menschliches Sprachverhalten in seiner immensen Komplexität zu simulieren und damit Rückschlüsse auf Intelligenz zuzulassen. Die Verarbeitung sprachlicher Symbole bleibt rein formal, eine Manipulation bedeutungsleerer Zeichen. Diesem Vorgang kann, so Steup, „im besten Fall ein versteinertes Weltbild derer zugeordnet werden, die diese Maschine realisiert haben.“⁵⁸ Erkennt also der Mensch in der Antwort des Systems eine sinnvolle Reaktion auf seine Eingabe, führt er den Sinn der Antwort selbst durch Projektion ein. An diesem Punkt greift der ELIZA-Effekt. Menschen neigen dazu, Gebrauchsgegenständen, speziell Maschinen, intendiertes, ja intelligentes Verhalten zu unterstellen.⁵⁹ Chatbot-Entwickler machen sich diesen Effekt zu Nutze.⁶⁰ Die Aussage, Chatbots verstünden auf menschenähnliche Weise Sprache, lässt sich nur aufrecht erhalten, wenn man den Menschen auf ein Reiz-Reaktions-System reduziert, wobei es letztlich unwesentlich ist, wie der Verstehensprozess aussieht, bzw. ob er überhaupt stattfindet, wenn der „Output“ stimmt.⁶¹

⁵⁸ Steup 1995:141

⁵⁹ Man führe sich dazu die Computermetaphorik und den alltäglichen Umgang mit dem PC vor Augen: „Er macht das nicht!“ „Jetzt zeigt er mir wieder diese Fehlermeldung!“

⁶⁰ So heißt es in der Dokumentation zur Demo-CD der Firma Kiwilogic: „Kinder streicheln zum Abschied das automatische Münzpfund vor dem Kaufhaus. [...] Selbst erwachsene Menschen weinen, wenn ihr Tamagotchi-Küken in den Cyberhimmel aufgefahren ist.“

⁶¹ In der deutschen ALICE Dokumentation wird die Frage gestellt, ob der Computer ein Reiz-Reaktions-System ist oder der Mensch. Interessant ist in dieser Hinsicht auch Searles „Chinesisches Zimmer“, ein Gedankenexperiment zur Bedeutung von Verstehen: Ein Mensch, der kein Chinesisch spricht, ist in einem Zimmer eingeschlossen. Unter der Tür hindurch werden ihm Zettel mit chinesischen Schriftzeichen gereicht, auf die er anhand einer Liste antworten muss. (Dieses Prinzip entspricht der Chatbot-Architektur). Die Antworten werden wieder unter der Tür hindurch gereicht, wo sie ein Chinese in Empfang nimmt. Die Antworten sind sinnvoll und in korrektem Chinesisch abgefasst. Hat der Mensch im Zimmer also verstanden? Ist dieser Punkt für den Chinesen überhaupt von Bedeutung? Searles Gedankenexperiment wird auch heute noch kontrovers diskutiert. s.a. Steup 1995

Die Einschätzung der Chatbots durch Vertreter der „konventionellen“ KI fällt pessimistisch aus. Menzel beispielsweise spricht von „völlig inadäquaten Surrogatlösungen“⁶² Und am Institut für Autonome Intelligente Systeme der Gesellschaft für Mathematik und Datenverarbeitung spricht man von „Intelligenzbetrugsprogrammen“⁶³ Intelligenz, so ein Argument, sei ohne körperlichen Bezug gar nicht möglich. Wer also Intelligenz simulieren wolle, der müsse einen Roboter bauen, der mitsamt seinem Körper autonom in der Welt handeln kann. Davon ist die Mensch-Maschine-Kommunikation über Chatbots, der jeglicher Kontextbezug fehlt, die auf jede parasprachliche Information verzichten muss, weit entfernt. Die etablierte KI-Forschung hält den Turing Test für den Beweis künstlicher (sprachlicher) Intelligenz für ungeeignet. Die Loebner-Wettbewerb Teilnehmer betrachteten den Test als Frage-Antwort-Spiel, das mit ein wenig „Lexikon und Grammatik“ funktioniere.⁶⁴

Der Loebner-Wettbewerb wird von den Teilnehmern tatsächlich eher spielerisch aufgefasst.⁶⁵ Anspruch ist, wie die oben angeführten Zitate von Wallace zeigen, weniger die Modellierung menschenähnlichen intelligenten Verhaltens als die perfekte Illusion. Es gibt zahlreiche Tipps und Tricks, mit denen die Jury getäuscht werden soll.⁶⁶ Dies reicht von „sei indirekt und mehrdeutig“ bis zu Listen mit den von Fremden am häufigsten gewählten Gesprächsthemen. Auf konversationsanalytische Ergebnisse aus der Linguistik wird dabei nicht zurückgegriffen, man verlässt sich eher auf statistische Ergebnisse (Zipf). Hier tritt die eher mechanische Sprachauffassung der Chatbot-Entwickler deutlich hervor. Dabei ist durchaus denkbar, dass die Ergebnisse konversationsanalytischer Forschung dazu beitragen könnten, die Illusion noch etwas weiter zu perfektionieren. Ob es jedoch möglich ist, Computern menschliche Sprachfähigkeit, den bewussten, selbstbestimmten Umgang mit Sprache zu „vermitteln“ bleibt fraglich.

⁶² Menzel 2000:644

⁶³ Lessmöllman 2000

⁶⁴ ebd.

⁶⁵ Auch Gewinner wie Hutchens stehen dem Contest als „Förderung der KI-Forschung“ durchaus kritisch gegenüber. (Hutchens 1997:18f)

⁶⁶ vgl. Whalen, Hutchens, Wallace und www.abenteuermedien.de/jabberwock/faq_de.html

Trotz aller Einwände, sind die Leistungen aktueller Chatbots verblüffend, vor allem, wenn man die relativ einfache Programmstruktur mit den komplexen Modellen der konventionellen KI vergleicht. Dennoch scheint sich ohne eben diesen Aufwand die Grenze zum wirklich menschlichen Sprachvermögen nie überschreiten zu lassen.

Literatur

- Görz, G.: Strukturanalyse natürlicher Sprache. - Bonn 1988
- Görz, G. (Hg.): Handbuch der künstlichen Intelligenz. - München 2000
- Hutchens, J.: How Hex works. - www.amristar.com.au/~hutch/he/how.html
- Hutchens, J.: How to pass the Turing Test by cheating. Online Dokument 1997
www.amristar.com.au/~hutch/
- Lessmöllmann, A.: Künstliche Intelligenz. Kumpel Computer. In: Die Zeit (41) 2000
- Menzel, W.: Sprachverarbeitung. Ein Überblick. In: Görz, G. (Hg.): Handbuch der künstlichen Intelligenz. - München 2000, S. 643-663
- Raphael, B.: The thinking computer. Mind inside matter. - San Francisco 1976
- Saygin, A. P.; Cicekli, I.; Akman, V.: Turing Test. 50 years later. -
- Schmid, K.: Die unermüdlichen Agenten. Sprachroboter arbeiten im Netz als Berater und Verkäufer. In: Die Zeit (13) 2001
- Steup, W.: Mensch und neue Maschinensysteme. Künstliche Intelligenz und Sprachverarbeitung in der politischen Bildung. - Schwalbach 1995
- Strube, G. et al: Kognition. In: Görz, G. (Hg.): Handbuch der künstlichen Intelligenz. - München 2000, S. 19-73
- Turing, A. M.: Computing machinery and intelligence. In: Mind (59) 1950, S. 433-460
- Vetter, M.: Softbots. Es menscht im Netz. In: PC Professional (11) 2001, S. 228-230
- Wallace, R.: Alice Dokumentation. - Mehrere Dateien unter www.alicebot.net
- Weizenbaum, J.: Computer power and human reason. From judgement to calculation. - San Francisco 1976
- Whalen, T.: Thom's Participation in the Loebner Competition 1995 or how I lost the contest and re-evaluated humanity. - <http://debra.dgbt.doc.ca/chat/story95.html>

Links (Auswahl)⁶⁷

Informationen über Chatbots / Hintergrundwissen

http://www.abenteuermedien.de/jabberwock/faq_de.html (Hintergrundwissen zur Chatbotprogrammierung)

<http://library.thinkquest.org/2705/> (Einführung in die künstliche Intelligenz)

<http://www.loebner.net/Prizef/loebner-prize.html> (Hintergrund zum Loebner Wettbewerb, Turing Test)

<http://bots.internet.com/> (allgem. zu Chatbots)

<http://www.botknowledge.com> (allgem. zu Chatbots)

<http://www.simonlaven.com> (die beste Seite für Einsteiger, zahlreiche Links und Papers)

<http://cogsci.ucsd.edu/~asaygin/tt/test.html> (Infos zum Turing Test)

Bots auf ALICE-Basis

<http://alicebot.org/live.html> (Auswahl von zahlreichen Alicebots)

<http://aimovie.warnerbros.com>

<http://www.german.alicebot.com>

<http://www.alicebot.net>

<http://elvis.alicebot.com>

<http://www.triumphpc.com/john-lennon/>

Weitere Spaßbots

<http://ecceliza.cjb.net/>

<http://www.tjhsst.edu/Psych/ch1/eliza.html>

<http://www.medical-tribune.de/000drelectric/index.html>

<http://www.parnasse.com/drwww.shtml> (Webdokter)

<http://www.robor.de>

<http://www.annette.ch.vu>

<http://elbot.de>

⁶⁷ Die folgenden Links waren bei Fertigstellung der Arbeit im Herbst 2001 aktuell.

Michael Storp: Chatbots. Möglichkeiten und Grenzen der maschinellen Verarbeitung natürlicher Sprache

<http://www.jackiestrike.com>

<telnet://debra.dgbt.doc.ca:3000>

<http://birch.eecs.lehigh.edu/alice/forbin.html> (Forbin lässt unterschiedliche Bots miteinander chatten)

<http://ciips.ee.uwa.edu.au/~hutch/hal/HEX/>

<http://ciips.ee.uwa.edu.au/~hutch/hal/HAL/>

<http://www.hamill.co.uk/mabel/>

<http://stick.us.itd.umich.edu/cgi-bin/chomsky.pl>

Kommerzielle Anbieter von Chatbotsoftware

<http://www.customer-care-solutions.de>

<http://www.kiwilogic.de>

<http://www.absolute.de>

<http://www.agentscape.de>

<http://artificial-life.com>

<http://www.novomind.com>

<http://www.piranhaz.com>

<http://virtualfriends.de>

Kommerzielle Chatbots

<http://www.agentland.com> (Übersicht über kommerzielle Bots)

<http://www.schweppes.de>

http://www.eye-track.de/welcome_e.html

<http://schwaebisch-hall.de/> (mit Sprachausgabe)

<http://www.appolinaris.de>