

Dennis Koch

„Menschliche“ Chatbots

Timo Reinhard
Fachbereich 3
Interaktivität
WS 06/07

Inhaltsverzeichnis

1. Einleitung.....	2
2. Der menschliche Dialog.....	2
3. Geschichte – Die Anfänge.....	4
4. ELIZA.....	5
1. Weizenbaum	
2. ELIZA Technik	
5. ALICE.....	8
1. Wallace	
2. ALICE Technik	
6. Vergleich der Technik.....	10
7. Resümee.....	11

Einleitung

Wenn man in letzter Zeit mal ein wenig durch das Internet surft, dann fällt einem auf, dass auf Firmenseiten hin und wieder statt den üblichen FAQ¹ eine nette Dame einen zum Chatten auffordert. In der Regel begrüßt sie einen mit den Worten „Hallo, mein Name ist Lisa. Wenn du Fragen zu unserem Produkt hast, kann ich sie dir beantworten. Schreib mir einfach!“. Nachdem man sich dann irritierter-, aber höflicherweise ebenfalls vorstellt, beginnt man also wie gefordert seiner neuen Bekannten Fragen zu dem Produkt zu stellen. Meistens stellt man dann aber im Verlauf des Gespräches fest, dass es sich bei Lisa leider um eine „Fachidiotin“ handelt, die selbst bei simplen Fragen nach dem Wetter auf einen kompetenten Mitarbeiter der Supportabteilung verweist. Das Gespräch endet dann häufig in unangenehmen Schweigen, bis man sich dazu entschließt sich freundlich zu verabschieden um weiter im Internet zu surfen.

Jedem, dem so eine Situation bekannt vorkommt, wird sich sicher daraufhin Gedanken über die Komplexität unserer Sprache gemacht haben.

Aus diesem Grund werde ich in dieser Hausarbeit zunächst auf den menschlichen Dialog und den Prozess des Erlernens und Verstehens der Sprache eingehen. Danach werde ich kurz die Anfänge der Chatbots erörtern um dann anhand der beiden Chatbots ELIZA und ALICE deren Technik zu analysieren und zu vergleichen. Abschließend werde ich einen Vergleich mit der menschlichen Sprache versuchen und die daraus resultierenden Konsequenzen aufzeigen.

Der menschliche Dialog

Um die Regeln eines menschlichen Dialogs zu erörtern, muss ich zunächst auf die dafür zugrunde liegende Sprache eingehen. Nach Gisela Szagun ist die Sprache in erster Linie ein Kommunikationsmittel, welches die folgenden Merkmale aufweist.

- Sie ist ein Symbolsystem. Wörter einer Sprache haben einen rein symbolischen Wert, da sie in der Regel, abgesehen von ein paar Lautmalereien, keinerlei Ähnlichkeit mit dem darzustellenden Objekt oder Ereignis besitzen.
- Sie ist kontextfrei. Das bedeutet, Sprache kann und wird meistens dafür benutzt über zeitlich oder räumlich entfernte Objekte oder Ereignisse zu sprechen. Sie ist nicht zwangsläufig mit der Gegenwart verknüpft und ist sogar mächtig genug uns unbekannte Dinge nur mit ihrer Hilfe zu beschreiben.

1 Frequently Asked Questions (dt. häufig gestellte Fragen)

- Sie wird kulturell vermittelt, ist also nicht angeboren und kann nur durch soziale Interaktion erlernt werden.
- Sie basiert auf einem kombinatorischen System insofern, dass sich ihre Symbole, durch Aneinanderreihung zu neuen Symbolen mit ähnlicher oder anderer Bedeutung ergänzen lassen. Ihr zugrunde liegt ein komplexes Regelwerk, die Grammatik. Mit ihr lassen sich aus einer begrenzten Anzahl von Wörtern unendlich viele neue Wörter und die daraus resultierenden Wortkombinationen und Sätze bilden. [1] S. 17-19

Aus diesen vier Punkten lässt sich bereits unschwer erkennen, dass das Erlernen einer Sprache nicht besonders trivial ist. So besteht die erste Hürde bereits bei der Verknüpfung von einfachen Worten und deren Bedeutungen, hinzu kommt noch die Schwierigkeit das selbe Objekt mit mehreren Worten zu kategorisieren. Beispielsweise ist jeder Rabe ein Vogel, aber nicht jeder Vogel ein Rabe. Das heißt Zuordnungen sind nicht unbedingt eindeutig und müssen daher beim Erlernen immer wieder angepasst und verfeinert werden. Ein Kleinkind sammelt bereits vor seinem Spracherwerb Eigenschaften von Objekten ihrer Umgebung, um diese dann anhand dessen zu kategorisieren. Im Laufe ihres Lernprozesses wird dann ein neuerlerntes (akustisches) Wort einem Objekt zugeordnet und ebenso auch auf Objekte der gleichen Kategorie ausgedehnt. Wenn dann weitere spezifische Eigenschaften des einzelnen Objektes erlernt worden sind, wird auch die Kategorie, welche dieses Wort beschreibt immer weiter eingeschränkt. Der Hund aus der Kategorie „Objekt auf vier Pfoten“ unterscheidet sich von einer Katze beispielsweise durch das Bellen, wodurch ihm alleine das Wort „Hund“ zugeordnet werden kann. Allerdings muss dieses Entscheidungskriterium ebenfalls erst erlernt bzw. erkannt werden, bevor es genutzt werden kann. [1] S. 132

Der andere wichtige Teil neben der Wortbedeutung ist die Grammatik. Diese unterteilt sich in Flexion und Syntaktik. Die Regelmäßigkeit dieser beiden Bereiche ist im hohen Maße sprachabhängig. Zudem kann im Deutschen die Flexion eines Wortes nur mit einer bestimmten Wahrscheinlichkeit richtig erfolgen. Das heißt es gibt keine festen Regeln nachdem jedes Wortes auf korrekte Weise gebeugt werden kann. Zum Beispiel kann der Numerus eines mit einem Konsonanten endenden Wortes durch anhängen eines „en“ geändert werden (vgl. das Hemd – die Hemden), allerdings ist das Erweitern mit „er“ genauso möglich (vgl. das Rind – die Rinder). Diese Unregelmäßigkeiten findet man in den anderen Wortarten genauso wieder und können zum Teil nur durch auswendiglernen korrekt gebildet werden.

Für das Erlernen der Syntaktik, schreibt Szagun, gibt es unter anderem den Ansatz der Universalgrammatik von Chomsky (1986). Allerdings stellen ihrer Meinung nach ihre empirischen

Daten diesen Ansatz in Frage. Nach Szagun erlernen Kinder die korrekte Syntaktik, indem sie anfangs neu erlernte Verben auf die lexikalisch gleiche Art und Weise wie sie sie kennengelernt haben wiederverwenden. Im Laufe der Zeit werden dann diese syntaktischen Schemata immer weiter abstrahiert, so dass diese dann nicht mehr von spezifischen Konstruktionstypen abhängig sind und von vornherein mit den korrekten Wortbeugungen benutzt werden. Die Syntaktik ist damit im Endeffekt eine Ansammlung von Regeln zur Aneinanderreihung von bestimmten Wortarten. [1] S.109 – 112

Damit also ein Dialog zwischen zwei menschlichen Gesprächspartnern geführt werden kann, sollten beide Gesprächspartner die erwähnten Aspekte einer Sprache so gut wie möglich beherrschen. Des Weiteren sollten beide Partner ähnliche Bedeutungen für die selben Wörter besitzen. Ein Dialog kommt dann dadurch zustande, dass jeder Gesprächsteilnehmer andere, aber für den Gegenüber nachvollziehbare Assoziationen zu dem aktuellen Gesprächsinhalt hat. Auf diese Weise kann eine Interaktion mit nahezu unendlichen Gesprächsthemen zwischen den Gesprächspartnern stattfinden.

Geschichte – Die Anfänge

Die Idee einer menschlichen Maschine ist nicht neu, bedenkt man die Andeutungen und Umsetzungen in vielen Science-Fiction Romanen wie „Star Wars“ oder „Per Anhalter durch die Galaxis“. Wer in diesem Zusammenhang als erstes eine menschlich sprechende Maschine erwähnt hat, lässt sich nur schwer ausmachen. Es lässt sich allerdings leicht herausfinden, wer als erstes konkrete Vorstellungen zum Nachweis menschlicher Eigenschaften eines Computers hatte.

Alan Mathison Turing, (* 1912 in London; † 1954 in Wilmslow) veröffentlichte 1950 seinen Turing-Test um die philosophische Frage zu klären, ob Computer denken können. [2]

Dieser Turing-Test besteht darin, dass eine Testperson nur mit Hilfe einer Tastatureingabe mit zwei unterschiedlichen Gesprächspartnern einen Dialog führt. Einer der beiden Partner ist menschlich, der andere ein Computer. Wenn die Testperson nach 5 Minuten noch nicht eindeutig sagen kann, wer von den beiden Partnern der Computer ist, gilt der Turing-Test als bestanden. Turing prophezeite, dass bis Ende des 20. Jahrhunderts Computersysteme existieren, bei denen 70% der beteiligten Testpersonen nicht sagen können, ob es sich um einen Menschen oder um einen Computer handelt. [3]

ELIZA

Joseph Weizenbaum

Joseph Weizenbaum wurde als zweiter Sohn des Ehepaars Jechied und Henriette Weizenbaum am 8. Januar 1923 in Berlin geboren. Da er wie seine Eltern Jude war musste er bereits 1936 aufgrund der nationalsozialistischen Rassengesetze des zweiten Weltkriegs Deutschland verlassen. Er emigrierte in die USA. Dort beendete er seine Schullaufbahn und absolvierte für ein Jahr ein Mathematikstudium in Detroit, bevor er dann bis zum Ende des zweiten Weltkriegs der US-Luftwaffe diente. Nach dem Krieg beendete er sein Studium und begann 1952 an einem Softwareprojekt der Bank of America mitzuwirken. 1963 gelangt er an das Massachusetts Institute of Technology (MIT) und forschte dort im Bereich Computernetzwerke. Im Laufe seiner Forschung entwickelte er 1966 das Computerprogramm ELIZA, welches dem Menschen die Möglichkeit gibt über die natürliche Sprache mit dem Computer zu kommunizieren. Ab 1996 lebte er wieder in Berlin und starb dort am 5. März 2008 an den Folgen eines Schlaganfalls. [4]

ELIZA (Technik)

ELIZA ist wie oben bereits erwähnt ein Programm, welches über die menschliche Sprache mit dem Benutzer kommunizieren kann. Es wurde 1966 von Joseph Weizenbaum veröffentlicht und sorgte zu dem Zeitpunkt für großes Aufsehen. Weizenbaum parodierte mit der als DOCTOR bekannten Version von ELIZA die Dialogstrategie eines Psychotherapeuten, da sie seiner Meinung nach besonders einfach zu realisieren war.

„Ein solcher Therapeut ist verhältnismässig leicht zu imitieren, da ein Grossteil seiner Technik darin besteht, den Patienten dadurch zum Sprechen zu bringen, dass diesem seine eigenen Äusserungen wie bei einem Echo zurückgegeben werden.“ [5]

Seiner Meinung nach wurde dieses Programm besonders bekannt und verbreitet, da es sich wie kaum ein anderes zu dieser Zeit dazu eignete die Fähigkeiten eines Computers in der Datenverarbeitung einem unspezialisierten Publikum zu demonstrieren. Allerdings machten ihn drei Dinge in diesem Zusammenhang nachdenklich. Erstens, dass einige praktizierenden Psychotherapeuten der Meinung waren dieses Programm ließe sich zu einer nahezu vollautomatischen Form der Psychotherapie ausbauen. Zweitens, die Geschwindigkeit mit der die Benutzer dieses Programms eine emotionale Beziehung zu dem Computer aufgrund der scheinbaren Menschlichkeit eingingen und drittens, die verbreitete Meinung dieses Programm sei die Lösung auf die Frage inwieweit ein Computer die natürliche Sprache verstehen kann.

Um diese Aspekte besser nachvollziehen zu können soll anhand eines englischen Beispieldialogs von Weizenbaum die Funktionsweise von Eliza erklärt werden [6].

Beispieldialog

```
1   Human: Men are all alike.
2   ELIZA: In what way?
3   Human: They're always bugging us about something or other.
4   ELIZA: Can you think of a specific example?
5   Human: Well, my boyfriend made me come here.
6   ELIZA: Your boyfriend made you come here
7   Human: He says I'm depressed much of the time.
8   ELIZA: I am sorry to hear you are depressed
9   Human: It's true. I am unhappy
10  ELIZA: Do you think coming here will help you not to be unhappy
11  Human: I need some help, that much seems certain.
12  ELIZA: What would it mean to you if you got some help
13  Human: Perhaps I could learn to get along with my mother.
14  ELIZA: Tell me more about your family
15  Human: My mother takes care of me.
16  ELIZA: Who else in your family takes care of you
17  Human: My father.
18  ELIZA: Your father
19  Human: You are like my father in some ways.
20  ELIZA: What resemblance do you see
21  Human: You are not very aggressive but I think you don't want me to notice
    that.
22  ELIZA: What makes you think I am not very aggressive
23  Human: You don't argue with me.
24  ELIZA: Why do you think I don't argue with you
25  Human: You are afraid of me.
26  ELIZA: Does it please you to believe I am afraid of you
27  Human: My father is afraid of everybody.
28  ELIZA: What else comes to your mind when you think of your father
29  Human: Bullies.
30  ELIZA: Does that have anything to do with the fact that your boyfriend
    made you come here
```

Wie man erkennen kann wirkt der Dialog auf den ersten Blick relativ menschlich. Erst auf den zweiten Blick scheinen manche Antworten etwas unüblich. Man beginnt an der „Menschlichkeit“ des Computerprogramms zu zweifeln und versucht daher das zugrundeliegende System zu entschlüsseln. Zwar ist dieses ohne das nötige Hintergrundwissen kaum möglich, doch fällt einem die Eintönigkeit mit der die Antworten erzeugt werden, recht schnell auf.

Das grundlegende Prinzip ist einfach zu erklären. Jeder eingegebene Satz wird nach bestimmten Schlüsselwörtern untersucht. Wenn ein Schlüsselwort gefunden wird, wird der eingegebene Satz mit Hilfe von Regeln, die diesem Schlüsselwort zugeordnet sind, umgestellt bzw. neu generiert. Wird kein Schlüsselwort gefunden, wird ein kontextloser Satz erzeugt.

Die internen Algorithmen, die auf der Suche nach dem richtigen Schlüsselwort sind, sind dafür umso komplexer. Das Programm muss mehrere Probleme lösen. Zum Einen muss aus dem

eingeebenen Text das Schlüsselwort herauskristallisiert werden, welches die größte Bedeutung für diese Nachricht hat. Da die Nachrichten von links nach rechts eingelesen werden, wird der Text insofern vereinfacht, dass Teilsätze in denen sich keine Schlüsselwörter befinden völlig fallen gelassen werden und ansonsten die höchste Priorität auf dem ersten gefunden Schlüsselwort liegt. Die Schlüsselwörter befinden sich in einer Liste der auch zugleich die Transformationsregeln zugeordnet ist.

Das nächste Problem, welches sich daraus ergibt, ist die Wahl der geeigneten Transformationsregel, falls mehrere Regeln diesem Schlüsselwort zugeordnet sind. Hier spielen dann bestimmte Masken eine Rolle mit deren Hilfe man auch einfache Kontexte, in denen diese Worte zu verstehen sind, erkennen kann. Das Schlüsselwort „You“ gefolgt von einem „are“ steht in einem anderem Kontext als „You“ gefolgt von einem Verb. Dementsprechend muss die Antwort auch anders gebildet werden. Somit kommen wir auch zu der Hauptaufgabe des Programms, der Bildung von Antworten.

Ein Programm kann die Bedeutung eines Wortes nicht verstehen. Weizenbaum vergleicht es mit einem Fremden, der nur sehr geringe Sprachkenntnisse besitzt [6], in diesem Fall die Schlüsselwörter und einige kontextfreien Phrasen. Wenn dieser Fremde nun den Satz „Ich bin heute sehr müde.“ hört, versteht er nur die Schlüsselwörter „Ich bin“. Der restliche Teil des Satzes wird erstmal isoliert. Zu den Schlüsselwörtern „Ich bin“ kennt er vielleicht eine Regel die besagt, dass alle Sätze, die mit diesem Ausdruck anfangen, in der Form umgewandelt werden können, dass die Wörter die hinter diesem Ausdruck stehen einfach durch voranstellen der Wörter „Warum bist du“ zu einem sinnvollen Fragesatz umformuliert werden können. Also: Warum bist du heute sehr müde?

Genauso macht es auch Eliza. Für das Programm ist es überhaupt nicht nötig zu wissen worüber gesprochen wird. Die Bedeutungen ist völlig irrelevant für die Konstruktion eines Antwortsatzes. Schwieriger wird es erst, wenn in dem Satz kein Schlüsselwort gefunden wird. In diesem Fall müssen Mechanismen greifen, die den Schein eines „intelligenten“ Dialogpartners erhalten können. Eliza greift da per Zufall auf vorgefertigte Phrasen oder bereits verwendete Sätze wieder zurück. Beispiele dafür sind die ersten beiden, sowie die letzte Antwort des Programms aus dem Dialog.

A.L.I.C.E.

Richard S. Wallace

Dr. Richard S. Wallace, der Erfinder von ALICE, ist 1960 in Portland, Maine geboren. 1989 machte er seinen Abschluss in Informatik an der Carnegie Mellon Universität und lebt jetzt zusammen mit

seinem Sohn und seiner Frau in San Francisco, Kalifornien. Von 1992 bis 1995 arbeitete er an der New York University unter anderem an einem Projekt zur Steuerung von Webcams mit natürlicher Sprache. Nachdem viele seiner Projekte aufgrund mangelnden Erfolgs eingestellt wurden, diagnostizierte man Mitte der 90er nach einem vereitelten Selbstmordversuch eine manische Depression. Wallace ist bis heute der Auffassung, dass seine Krankheit der Grund für seine Kündigung war und hat daher immer wieder Auseinandersetzungen mit Universitäten und Professoren.

Als er 1995 arbeitslos wurde begann er die Arbeit an dem Programm ALICE. [7] Ursprünglich als SETL Programm geschrieben portierte er es 1998 auf Java und veröffentlichte es unter der GNU GP License. Seitdem arbeiten mehr als 500 Programmierer weltweit an diesem Projekt mit. Wallace ist nun Vorstandsvorsitzender und Mitbegründer der ALICE Artificial Intelligence Foundation. [8] Mit ALICE gewann er 2000, 2001 und 2004 den Loebner Preis. Ein Preis bei dem die menschlichsten Chatbots mit Hilfe des Turing-Tests ausgezeichnet werden. [9]

A.L.I.C.E (Technik)

A.L.I.C.E steht für Artificial Linguistic Internet Computer Entity. Es basiert auf AIML einer Markupsprache um ähnlich wie bei ELIZA Regeln und Antworten für bestimmte Schlüsselwörter bereitzustellen. Sie fungiert als eine Art Datenbank, in denen die möglichen Antworten auf bestimmte Schlüsselwörter gespeichert sind.

```
1 Human: How are you?
2 ALICE: I am fine, thank you.
3 Human: What are your hobbies?
4 ALICE: I like robots, computers, and chatting online.
5 Human: Why do you like robots?
6 ALICE: I am a robot. We robots stick together.
7 Human: That's great.
8 ALICE: Thanks, unknown person.
9 Human: I am sorry, my name is Peter.
0 ALICE: There is no need to apologize. Hi there Peter.
11 Human: Do you ever go outside?
12 ALICE: I go wherever my botmaster takes me.
13 Human: Who is your botmaster?
14 ALICE: I was created by Dr. Richard S. Wallace.
15 Human: I think I know him.
16 ALICE: Why? What do you think of him?
17 Human: I met him only once.
18 ALICE: What did you do with him when you first met?
19 Human: We talked about the weather.
20 ALICE: Of course I remember you well Peter. We were talking about only.
21 Human: What do you think about the weather?
22 ALICE: The general condition of the atmosphere at a particular time and
place, with regard to the temperature, moisture, cloudiness, etc.
23 Human: It is time to go now.
```

24 ALICE: Oh. It is about that time now.
25 Human: Goodbye.
26 ALICE: Bye for now.

Auch hier bemerkt man schnell, dass manche Antworten nicht besonders passend sind, allerdings erscheinen sie nicht so „mechanisch“ konstruiert. Ebenfalls auffällig ist auch, dass bei diesem Chatbot der Nutzer nicht immer auf der Seite des Antwortenden steht. Dadurch wirkt der Dialog nicht so einseitig und bekommt menschlichere Züge. Die Frage ist nun auf welche Weise ALICE programmiert wurde. Das grundlegende System ist auch hier sehr einfach. Nachdem der Satz in seine Bestandteile zerlegt wurde, wird in einer AIML-Datei nach den entsprechenden Schlüsselwörtern gesucht. Wird keine Maske gefunden auf die der eingegebene Satz passt, wird ebenfalls auf eine Menge vordefinierter Ausweichsätze zurückgegriffen. Um das System besser zu verstehen werden hier drei beispielhafte Ausschnitte einer deutschen AIML-Datei beschrieben.

Ausschnitt 1:

```
<category>
  <pattern>HAST DU * GESEHEN</pattern>
  <template>Tut mir leid, ich habe <star /> nicht gesehen. Ich kann zur Zeit
    gar nichts <set name="thema">sehen</set>. Ich habe noch keine Augen.
  </template>
</category>
```

Ausschnitt 2:

```
<category>
  <pattern>WORÜBER REDEN WIR GERADE</pattern>
  <template>Wir reden gerade ueber <get name="thema"/>.</template>
</category>
```

Ausschnitt 3:

```
<category>
  <pattern>ALLE *</pattern>
  <template><random>
    <li>Ich glaube, es gibt auch Ausnahmen.</li>
    <li>Nicht wirklich alle, oder?</li>
    <li>Das ist jetzt aber stark verallgemeinert.</li>
  </random></template>
</category>
```

Der eingegebene Satz wird mit den innerhalb der pattern-Tags eingeschlossenen Maske verglichen. Das bedeutet, es wird zum Beispiel im ersten Abschnitt überprüft ob sich die Worte HAST, DU und GESEHEN in dieser Reihenfolge innerhalb des Eingabesatzes befinden. Das Sternchen ersetzt eine unbegrenzte Anzahl unbekannter Wörter. Diese Wörter werden aber nicht verworfen sondern in einer temporären Variable zwischengespeichert, da sie in einer möglichen Antwort nochmal gebraucht werden könnten. Die Antwort auf dieses Pattern wird von dem template-Tag eingeschlossen. Innerhalb dieses Tags können neben dem reinen Text auch, wie in diesem Beispiel, zusätzliche Anweisungen stehen. Der XML-Ausdruck <star/> wird, wie oben bereits angedeutet,

durch die gespeicherte, temporäre Variable ersetzt. Aufgrund dieses Bezuges scheint das Programm ähnlich wie bei ELIZA verstanden zu haben worum es geht. Im Verlauf der Beispielantwort wird aber auch noch eine weitere Anweisung aufgerufen. Der Ausdruck `<set name="thema">sehen</set>` erzeugt eine Variable „thema“ mit dem, durch die Tags eingeschlossenen, Wert. Auf diese Weise können Variablen über einen längeren Zeitraum gespeichert werden und durch eine Anweisung, wie im zweiten Ausschnitt, auch wieder aufgerufen werden. Der Beispieldialog nutzt diese Technik um den in Zeile 9 eingegebenen Namen in Zeile 20 wieder zu erwähnen. In Zeile 8 war dem Programm der Name noch nicht bekannt, weswegen die Antwort mit dem vorinitialisierten Wert „unknown person“ generiert wurde.

Die ersten beiden Ausschnitte besitzen für einen bestimmten Eingabesatz nur eine vorgeschriebene Antwortmöglichkeit. Der dritte Ausschnitt zeigt, dass das nicht zwangsläufig so sein muss. Auf einen Eingabesatz beginnend mit dem Wort „Alle“ kann per Zufallsgenerator einer der Einträge innerhalb des random-Tags als Antwort zurückgegeben werden. Für den Fall, dass keine Maske für den eingegebenen Satz existiert, wird ähnlich wie in diesem Ausschnitt auf eine Liste kontextfreier Sätze zurückgegriffen.

Vergleich der Technik

Vergleicht man ELIZA und ALICE miteinander, erkennt man, dass beide Systeme auf ziemlich ähnliche Art und Weise funktionieren. Sie zerlegen die eingegebenen Sätze um darauf basierend eine Antwort zu erzeugen. Der Unterschied besteht lediglich darin, wie die Antwort erzeugt wird. ELIZA nutzt größtenteils nur Elemente des gerade eingegebenen Satzes wohingegen ALICE auf eine Datenbank mit vordefinierten Antwortmöglichkeiten zurückgreift. Zusätzlich speichert sich ALICE für den Dialog wichtige Variablen. Durch die Einschränkung von ELIZA, nicht auf eine Datenbank zurückgreifen zu können, gibt es nur die Möglichkeit mit Hilfe von sehr einfachen Regeln eingegebene Sätze in Fragen umzuwandeln. Zwar wird auf diese Weise der Dialog im Fluss gehalten, doch entsteht schnell eine Eintönigkeit und der Benutzer entlarvt es als Computerprogramm. ALICES Technik ist in diesem Punkt weiter ausgereift, kann es aber nur mit Hilfe einer umfangreichen Datenbank an Antworten realisieren. Um die Komplexität dieser Datenbank etwas zu reduzieren, ist sie mit sehr einfachen austauschbaren Variablen ausgestattet.

Beide Systeme versuchen aber nur einen menschlichen Dialog zu imitieren. ELIZA reduziert die menschliche Sprache auf einfache grammatische Regeln und ALICE reduziert den Dialog mit einer großen Menge an Entscheidungsvariablen auf eine typische Anzahl Wörter und Sätze. Sie sind nicht

in der Lage Assoziationen mit den Bedeutungen der Wörter zu verbinden, denn dieses würde eine dynamische Kategorisierung der Bedeutung von Wörtern, wie im Lernprozess von Gisela Szagun beschrieben, erfordern. Des Weiteren sind beide Systeme nicht in der Lage grammatikalische Konstruktionen im Sinne der Flexion zu erzeugen. Sie sind allein auf die vordefinierten Parameter eingeschränkt. Diese Einschränkung resultiert aus der schon im ersten Kapitel erwähnten unbestimmten Wahrscheinlichkeit, Wörter, vor allem in der deutschen Sprache, richtig zu beugen.

Nichtsdestotrotz soll hier erwähnt werden, dass neben den Schwachstellen in der technischen Realisierung, einige Merkmale auffallen, die einen dazu verleiten seinem Chatpartner kurzzeitig Persönlichkeit zuzuschreiben. Vor allem das Programm ALICE suggeriert eine gewisse Intelligenz, da die Antworten aus Sicht des Benutzers in den meisten Fällen in den Kontext passen. Kleinere Fehler, wie ein nicht initialisierter Namen im Beispieldialog Zeile 8 – 10, fallen einem unter Umständen gar nicht mehr auf. Man wird sogar dazu verleitet sich zu entschuldigen. Wäre man sich bewusst, dass es sich um einen Computerprogramm handelt, käme man sicherlich nicht so leicht auf die Idee. Wallace hat unter anderem den Erfolg seines Chatbots durch die stetig steigende Konversationslänge realisiert. Hätte der Chatbot nicht den Schein eines intelligenten Dialogpartners gewahrt, wären die Tests auf seiner Website nur auf wenige Wechselbeziehungen begrenzt gewesen. [7]

Resümee

Die hier genannten Chatbots vereinfachen die menschliche Sprache auf das Nötigste und imitieren auf diese Weise einen intelligenten Dialog. Beindruckend ist meiner Meinung nach aber die Tatsache, dass es in vielen Fällen funktioniert. Technisch gesehen, ist so ein Chatbot kaum mehr als ein paar geschickt verschachtelte Entscheidungsanweisungen. Betrachtet man ihn allerdings aus der Sicht eines unwissenden Benutzers, kann dieses Programm schnell menschliche Züge erhalten. Das Hauptproblem besteht darin die Antworten der Chatbots möglichst allgemein auf alle Themen anzupassen. Dieses Problem lässt sich nur auf zwei Wege lösen. Entweder man schränkt die Themenwahl des Benutzers ein, oder man vervollständigt die Datenbank auf alle möglichen Themen. Die Einschränkung der Themenwahl lässt sich relativ einfach realisieren, indem man den Chatbot in einen Kontext stellt, welcher nur ein Thema zulässt. So bietet sich beispielsweise, die schon in der Einleitung erwähnte Anwendung als FAQ-Chatbot oder Verkaufsberater an. Eine vollständige Datenbank zu allen Themen zu schaffen scheint im Moment etwas utopisch, könnte aber im Zusammenhang mit einem möglichen zukünftigen, semantischen Web in eine realistische Nähe rücken.

Literaturverzeichnis

- [1] Szagun, G. (2006). Sprachentwicklung beim Kind: ein Lehrbuch. Weinheim: Beltz.
- [2] Hodges, A. The Alan Turing Home Page,
<http://www.turing.org.uk/turing/>
- [3] Turing, A. M. (1950). Computing machinery and intelligence.
<http://www.loebner.net/Prizef/TuringArticle.html>
- [4] Löw, W. (2001) WWW = Wissen Werte Weizenbaum - aus neurobiologischer Sicht.
http://www.meduniwien.ac.at/agmb/mbi/2003_1/41-43loew.pdf
- [5] Gymnasium Laurentianum, Joseph Weizenbaum und Eliza
<http://www.laurentianum.de/leinfo06.htm#kap3>
- [6] Weizenbaum, J. (1966). ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine
<http://i5.nyu.edu/~mm64/x52.9265/january1966.html>
- [7] Thompson, C. (2002). Approximating Life
[http://query.nytimes.com/gst/fullpage.html?
res=9E02E2D61F3EF934A35754C0A9649C8B63&sec=&spon=&partner=permalink&exprod=per
malink](http://query.nytimes.com/gst/fullpage.html?res=9E02E2D61F3EF934A35754C0A9649C8B63&sec=&spon=&partner=permalink&exprod=permalink)
- [8] Wallace, S. R. Biografie
<http://www.alicebot.org/bios/richardwallace.html>
- [9] Home Page of The Loebner Prize in Artificial Intelligence
<http://www.loebner.net/Prizef/loebner-prize.html>

Alle Links wurden am 16.9.2008 auf Gültigkeit überprüft.