

Künstliches Themenbewusstsein in natürlichen Dialogen

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

Alexa Breuing

Künstliches Themenbewusstsein in natürlichen Dialogen

Alexa Breuing
AG Wissensbasierte Systeme (Künstliche Intelligenz)
Technische Fakultät
Universität Bielefeld
Postfach 10 01 31
D-33501 Bielefeld
Germany
E-Mail: abreuing@techfak.uni-bielefeld.de

Von der Technischen Fakultät der Universität Bielefeld zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Dissertation.

Dekan der Fakultät: Prof. Dr. Jens Stoye
Erster Gutachter: Prof. Dr. Ipke Wachsmuth
Zweiter Gutachter: Prof. Dr. Ulrich Furbach

Einreichung der Arbeit: 23. Oktober 2012
Tag der Disputation: 30. Januar 2013

Die offizielle Druckversion wurde auf alterungsbeständigem Papier gemäß DIN-ISO 9706 gedruckt.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich auf dem Weg zur Promotion begleitet und unterstützt haben. Besonderer Dank gilt meinem Doktorvater, Herrn Prof. Dr. Ipke Wachsmuth, der bereits während meines Informatikstudiums an der Universität Bielefeld meine Begeisterung für die Mensch-Maschine-Kommunikation und für die wissenschaftliche Arbeit mit dem virtuellen Agenten Max weckte. Die Möglichkeit, unter seiner Obhut zu promovieren, bot mir nicht nur eine ausgezeichnete und umfangreiche wissenschaftliche Betreuung, sondern förderte gleichermaßen meine fachliche, berufliche sowie persönliche Weiterentwicklung. Für die bereitwillige Übernahme des Zweitgutachtens bedanke ich mich herzlich bei Herrn Prof. Dr. Ulrich Furbach.

Ich danke der gesamten Arbeitsgruppe Wissensbasierte Systeme, die mir stets eine Umgebung und Atmosphäre bot, die mir die Arbeit leicht fallen ließ. Insbesondere der lebendige Austausch mit meinen lieben Gleichgesinnten Hana, Nhung und Maha gab mir viel Rückhalt, Fröhlichkeit und ein beruhigendes Gefühl des Zusammenhalts. Ich danke Kevin für die ungemein wertvolle praktische Unterstützung, meinen Projektpartnern Ulli und Alexandra für fachliche Ratschläge und konstruktive Diskussionen und Cord für sein beständiges Interesse an meinem Thema.

Meiner Familie und meinen Freunden bin ich für die Rücksichtnahme, ihren großartigen Aufbauarbeiten und das permanente Daumendrücken über die ganzen Jahre hinweg dankbar. Der größte Dank gilt Christian, für seinen Glauben an mein Können und seine liebevolle Aufmerksamkeit.

Inhaltsverzeichnis

Abbildungsverzeichnis	v
Tabellenverzeichnis	vii
1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung und Aufgabenbeschreibung	3
1.3 Das KnowCIT-Projekt	7
1.4 Aufbau der Arbeit	8
2 Themen in Dialogen	11
2.1 Sprachbasierte Kommunikation	11
2.1.1 Dialog	12
2.2 Themen in sprachlichen Diskursen	15
2.2.1 Der Begriff „Thema“	15
2.2.2 Themenbewusstsein	19
2.3 Das mentale Lexikon	29
2.4 Zusammenfassung	34
3 Grundlagen künstlichen Themenbewusstseins	35
3.1 Konzeptwissen für Maschinen	36
3.1.1 Wissensrepräsentation	37
3.1.2 Wikipedia	42
3.1.3 Diskussion	44
3.1.4 Indexierung von Textinhalten	46
3.2 Automatische Themenerkennung in Texten	49
3.2.1 Textähnlichkeit	49
3.2.2 Textkategorisierung	55

3.2.3	Das TDT-Projekt	60
3.3	Themenbewusstsein in der konversationalen Mensch-Maschine-Interaktion	62
3.3.1	Konversationale Schnittstellen	62
3.3.2	Beispielsysteme für künstliches Themenbewusstsein	66
3.4	Zusammenfassung	71
4	Künstliches Themenbewusstsein für den konversationalen Agenten Max	75
4.1	Aufgabenbeschreibung	75
4.2	Das Gesamtkonzept	76
4.3	Das Themenmodell	77
4.3.1	Arbeitsdefinition „Dialogthema“	78
4.3.2	Der Themenerkenner	78
4.3.3	Themeninformationen	82
4.4	Integration der Themeninformationen	85
4.4.1	Anforderungsanalyse	86
4.4.2	Einbettung des Themenerkenners in die Systemarchitektur	86
4.4.3	Themenbasierte Anpassung des konversationalen Verhaltens	88
4.5	Zusammenfassung	90
5	Ein Modell zur Erkennung aktueller Dialogthemen	93
5.1	Wikipedia als Quelle konzeptueller Wissensstrukturen	94
5.1.1	Informationsspezifikation	95
5.1.2	Informationsindexierung	97
5.2	Dynamische, online-fähige und globale Themenerkennung	99
5.3	Wikipedia-basierte Umsetzung des Themenerkenners	102
5.3.1	Identifikation der Konzeptthemen	104
5.3.2	Identifikation des Dialogthemas	111
5.3.3	Algorithmus	114
5.3.4	Ableitung der Themeninformationen	117
5.4	Evaluation	124
5.4.1	Bewertungsmaße	124
5.4.2	Vorbereitende Evaluation auf Basis von Zeitungsartikeln	125
5.4.3	Evaluation auf Basis des CUBE-G-Korpus	132
5.5	Zusammenfassung	137

6	Integration des Themenmodells	139
6.1	Der konversationale Agent Max	140
6.1.1	Dialogfähigkeit	141
6.2	Integration des Themenerkenners in die Systemarchitektur	142
6.3	Integration der Themeninformationen in das Konversationsverhalten . .	145
6.3.1	Emulierte Fähigkeit zur Themenbenennung	145
6.3.2	Emuliertes Bewusstsein für Themeneinführungen	146
6.3.3	Emuliertes Bewusstsein für Themenfortführungen	148
6.3.4	Emuliertes Bewusstsein für Themenwechsel	148
6.3.5	Emuliertes Bewusstsein für Themenwiedereinführung	151
6.4	Anwendungsbeispiel	151
6.4.1	Dialogszenario	152
6.4.2	Beispieldialog	154
6.5	Zusammenfassung	157
7	Resümee	159
7.1	Ergebnisse	159
7.2	Fazit	163
7.3	Ausblick	164
	Literaturverzeichnis	167

Abbildungsverzeichnis

1.1	Realisierung künstlichen Themenbewusstseins für konversationale Schnittstellen	4
2.1	Kommunikationsarten	12
2.2	Der Dialog als Gesprächsform zwischen zwei Personen	13
2.3	Elemente eines Dialogs	14
2.4	Lexikalische Mehrdeutigkeiten	32
3.1	Beispiel für ein semantisches Netz	40
3.2	Einzelne Bestandteile eines Wikipedia-Artikels	43
3.3	Beispiel für einen Vektorraum	51
3.4	Vektorraum mit Dokumentvektoren	53
3.5	Kategoriensystem der Wikipedia	57
3.6	Eine Schleife innerhalb des Kategorienbaums	57
3.7	Konzeptualisierung und Kategorisierung auf Basis von Wikipedia	59
3.8	Beispiele konversationaler Agenten	65
4.1	Gesamtarchitektur zur Realisierung künstlichen Themenbewusstseins	76
4.2	Architektur des dynamischen, online-fähigen und globalen Themenerkenners	82
4.3	Bestehende Architektur des Dialogsystems von Max	87
4.4	Integration der Themeninformationen in die Deliberative Komponente	88
5.1	Gesamtarchitektur des Themenmodells	94
5.2	Wikipedia-Artikel über den virtuellen Agenten Max	96
5.3	Basisprozesse zur Identifikation von Dialogthemen	100
5.4	Konzeptzusammenhänge innerhalb des Wikipedia-Kategoriengraphen	102
5.5	Gesamtprozess der automatischen Themenerkennung	103
5.6	Konzeptualisierung einer Äußerung	107

5.7	Vektordarstellung eines Konzepts	108
5.8	Kategorisierung von Konzepten	110
5.9	Vektordarstellung aller Konzepte einer Äußerung	111
5.10	Berechnung des Dialogthemenvektors	113
5.11	Graphische Darstellung der Themeninformationen	117
5.12	Identifikation des Dialogthemas für zwei Äußerungen	119
5.13	Erkennung eines möglichen und durchgeführten Themensprungs	121
5.14	Beispiel für eine Themenwiedereinführung	123
5.15	Precision und Recall	125
5.16	Auszug aus der Online-Studie	135
6.1	Einbettung des automatischen Themenerkenners in die Architektur des konversationalen Agenten Max	140
6.2	Max in der Rolle als Museumsführer	141
6.3	Darstellung des Themenagenten	143
6.4	Kommunikation zwischen Themen- und Dialogagent	144
6.5	Beispieldialog	156

Tabellenverzeichnis

3.1	Quellen konzeptuellen Wissens im Vergleich	45
3.2	Beispiel für einen Index	46
3.3	Häufigkeiten für die Gewichtung von Termen in Dokumenten	52
4.1	Überblick über die Themeninformationen	85
5.1	Auszug aus dem Artikelindex	99
5.2	Auszug aus dem Kategorienindex	99
5.3	Evaluationsergebnisse der automatischen Themenidentifikation und Themenbenennung	127
5.4	Evaluationsergebnisse des globalen Themenerkenners	130
5.5	Schwellwert-Bestimmung	131
5.6	Themen der CUBE-G-Dialoge und ihre Kategorienrepräsentationen	133
5.7	Ergebnisse der Online-Studie	136
6.1	Textbasierte Spezifikation adäquater Dialogthemen	147
6.2	Liste der vordefinierten Dialogthemen für das Beispielszenario	155

1 Einleitung

Taking part in a conversation is not an obviously hard task to the participants. On the surface it seems no more difficult than reading a story. Indeed, it would seem to be easier. Children conduct conversations before they can read [...]. Nonetheless, [...] conversation is one of our most challenging problems. (Schank, 1977, S. 421)

Gespräche zwischen Menschen sind gekennzeichnet durch kohärente Abfolgen sprachlicher Äußerungen, welche der gesamten Konversation eine thematische Struktur verleihen. Grundlage für diese Kohärenz und für das Gelingen einer sowohl bedeutungsvollen als auch flexiblen Konversation liefert die menschliche Fähigkeit, sich eines Gesprächsthemas *bewusst* zu sein. Mit dem Ziel, thematisch kohärente Gespräche auch zwischen Menschen und interaktiven Systemen zu ermöglichen, werden im Rahmen der vorliegenden Arbeit die Automatisierung menschlichen Themenbewusstseins und die Übertragung dieses künstlichen Bewusstseins auf einen virtuellen Agenten unternommen.

1.1 Motivation

War in der Vergangenheit noch eine Anpassung seitens des Benutzers erforderlich, um ein Computersystem effizient zu bedienen, wird heutzutage intensiv daran geforscht, eine natürliche und dadurch dem Menschen angepasste Interaktion mit Maschinen zu ermöglichen. Ein Beispiel dafür sind sprachbasierte Mensch-Maschine-Schnittstellen, bei denen Anwender mittels gesprochener oder textueller, natürlichsprachlicher Eingaben mit der Maschine kommunizieren können. Statt maschinennaher Kommandos, deren Syntax durch die jeweils verwendete Programmiersprache festgelegt wird, dienen intuitiv formulierte Äußerungen als Anweisungen an das System. Eine Weiterentwicklung solch sprachbasierter Benutzerschnittstellen sind intelligente, konversationale Agenten, von Cassell et al. (2000) als sogenannte *Embodied Conversational Agents*

(ECAs) bezeichnet. ECAs, im Folgenden *konversationale Agenten* genannt, sind virtuell dargestellte, autonome Charaktere, deren Modalitäten denen der menschlichen Konversation entsprechen. Das heißt, sie können Sprache, Gestik und Mimik von Menschen sowohl verstehen als auch produzieren und schaffen damit die Basis für eine auf natürlicher Konversation basierende Mensch-Maschine-Interaktion.

Zwischenmenschliche Kommunikation beruht jedoch nicht nur auf dem Sprechvermögen und dem Sprachverständnis der beteiligten Personen an sich. Gespräche zwischen Menschen zeichnen sich insbesondere durch kohärente Abfolgen sprachlicher Äußerungen aus, die es den Sprechern erlauben, sich über bestimmte Gesprächsinhalte auszutauschen. Wesentlich ist hier der Aspekt des *Austausches*, denn welche Themen im einzelnen während eines Gesprächs adressiert werden, hängt von den Beiträgen beider Gesprächsteilnehmer ab. Der Themenverlauf der gesamten Konversation etabliert sich daher interaktiv und ist das Ergebnis eines *joint project* der beteiligten Dialogpartner (Clark, 1996). Grundlage für diesen wechselseitigen Prozess liefert die menschliche Fähigkeit, sich des Themas eines Gesprächs *bewusst* zu sein. Mit diesem Bewusstsein geht sowohl die Wahrnehmung als auch die Initiierung von Aktionen einher, die den thematischen Ablauf eines Gesprächs beeinflussen. Dazu gehört beispielsweise die Erkennung von Themenvorschlägen oder Themenwechseln.

Virtuellen Charakteren, welche die zuvor beschriebenen Eigenschaften eines konversationalen Agenten aufweisen, könnte aufgrund ihres menschenähnlichen Erscheinungsbilds instinktiv ein solches Themenbewusstsein zugeschrieben werden. Demnach erwarten ihre menschlichen Kommunikationspartner, dass der Agent zum einen weiß, worüber sie sich gerade mit ihm unterhalten, zum anderen aber auch, dass das zugrundeliegende System dieses Wissen nutzt, um den Agenten *adäquat* auf ihre Äußerungen reagieren zu lassen. Das Wissen über das aktuelle Gesprächsthema hilft uns Menschen zum Beispiel, sprachliche Mehrdeutigkeiten und Referenzen aufzulösen. Eine Frage wie „*Wann wurde Moore geboren?*“ lässt sich eindeutig beantworten, wenn zum Zeitpunkt der Fragestellung über James Bond gesprochen wurde und der mehrdeutige Begriff „Moore“ somit als Referenz auf den Schauspieler Roger Moore interpretiert werden kann. Agenten mit Frage-Antwort-Fähigkeiten könnten ähnlich verfahren und den Suchraum für eine eingegebene Frage aufgrund dieser Kontextinformation thematisch einschränken.

Adäquat bedeutet aber auch, dass die Antworten des Agenten bestenfalls themati-

sche Überlappungen zu den vorherigen Äußerungen aufweisen und so ein inhaltlicher Zusammenhang zwischen den einzelnen Gesprächsbeiträgen hergestellt wird. Auf diese Weise gelingt es dem Agenten, mit seinem menschlichen Gegenüber ein gemeinsames Gesprächsthema aufzubauen und bedeutungsvolle, thematische Konversationen zu führen. Dies garantiert eine für den Menschen intuitive Interaktionsform und führt zu einer erhöhten Akzeptanz und Glaubwürdigkeit (*believability* (Cassell, Bickmore, Campbell, Vilhjálmsón & Yan, 2000)) des konversationalen Agenten als gleichwertigen Konversationspartner.

Im Bereich der Mensch-Maschine-Interaktion gibt es erst wenige Bestrebungen, ein automatisiertes Themenbewusstsein für den künstlichen Interaktionsteilnehmer zu realisieren. Bestehende Ansätze bestimmen die Themen anhand einzelner Äußerungen, wobei die Wechselbeziehung hintereinander auftretender Gesprächsbeiträge außer Acht gelassen wird. Mit der von Cassell et al. (2000) festgelegten Anforderung, sich bei der Implementierung eines konversationalen Agenten und seiner zugrunde liegenden Architektur stets an Eigenschaften realer, interpersoneller Konversation zu orientieren, wird die im Rahmen dieser Arbeit durchgeführte Realisierung künstlichen Themenbewusstseins stark von dem sequentiellen und interaktiven Charakter zwischenmenschlicher Gespräche bestimmt. Demzufolge leistet diese Arbeit einen Beitrag zur Verbesserung der Interaktion zwischen Menschen und sprachbasierten Systemen durch die Realisierung kohärenter, thematischer Gespräche zwischen Mensch und konversationalem Agent und schafft gleichzeitig eine Grundlage für die Erschließung neuer Erkenntnisse bezüglich menschlicher konversationaler Verhaltensweisen.

1.2 Zielsetzung und Aufgabenbeschreibung

Ziel dieser Arbeit ist die Entwicklung künstlichen Themenbewusstseins, welches, wie in Abbildung 1.1 dargestellt, die Erkennung und Analyse von Gesprächsthemen in laufenden, natürlichen Konversationen und die Integration der daraus resultierenden Themeninformationen in das Dialogsystem eines konversationalen Agenten beinhaltet. Sowohl Konzeption als auch Implementation sind dabei so auszuführen, dass die Simulation menschlichen Themenbewusstseins gelingt und die kommunikativen Fähigkeiten des Agenten mittels dieses Bewusstseins erweitert werden. Mit Blick darauf, dass hier also eine Kompetenz simuliert werden soll, die menschlichen Fähigkeiten sehr nahe

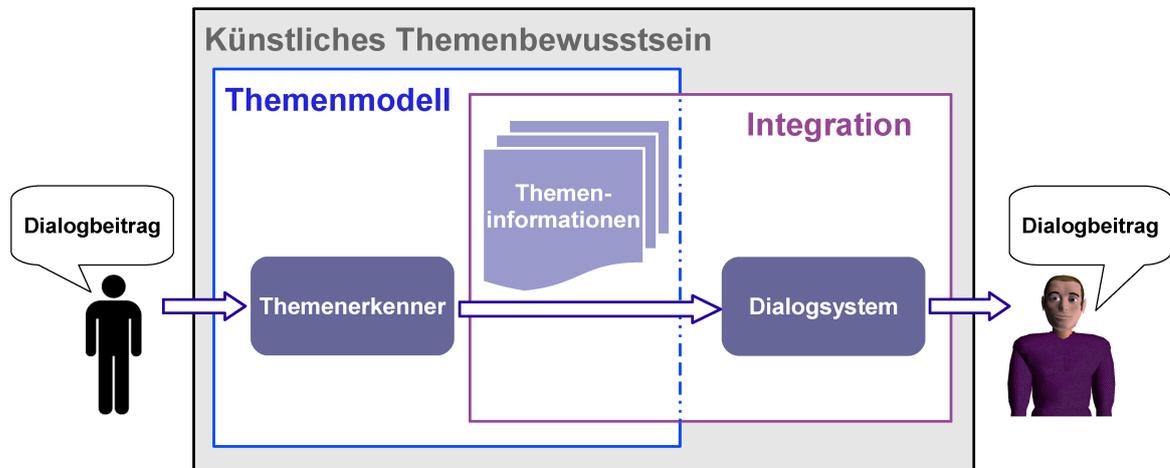


Abbildung 1.1: Überblick über die Gesamtarchitektur zur Realisierung künstlichen Themenbewusstseins für konversationale Schnittstellen innerhalb der natürlichen Mensch-Maschine-Interaktion.

kommt, wird im weiteren Verlauf der Arbeit von der *Emulation* menschlichen Themenbewusstseins gesprochen. Auf Grundlage dieser Emulation wird eine Schnittstelle geschaffen, die kohärente und somit natürliche Gespräche zwischen Menschen und Agenten erlaubt und gleichzeitig eine Plattform für die Erforschung weiterer Aspekte menschlichen Konversationsverhaltens bereitstellt.

Während einer Konversation werden unterschiedliche Informationen zur Erkennung des aktuellen Themas herangezogen. Neben den vorliegenden Gesprächsbeiträgen der teilnehmenden Konversationspartner helfen zusätzliche Kontextinformationen, wie beispielsweise die umgebende Gesprächssituation oder Kenntnisse aus der letzten gemeinsamen Unterhaltung, den möglichen thematischen Rahmen einzuschränken und die Themenerkennung durchzuführen. Innerhalb des Themenmodells werden diese Informationen nicht in den Prozess der automatisierten Themenerkennung einbezogen, dessen Fokus ausschließlich auf der Verarbeitung der sprachlichen Beiträge liegt. Allein auf Basis dieser Inhalte soll dem künstlichen Konversationspartner ermöglicht werden, angesprochene Themen zu identifizieren und sie zu benennen, sowie thematische Wechsel innerhalb einer Unterhaltung zu erkennen. Bei der Integration der ermittelten Themeninformationen in das konversationale Verhalten des künstlichen Zielsystems spielen diese Kontextinformationen jedoch eine große Rolle, da die Wahl eines Gesprächsthemas oftmals durch die gegebene Situation beeinflusst wird.

Anhand der Zielsetzung lässt sich die Arbeit in zwei Aufgabenschwerpunkte unterteilen, die im Folgenden kurz vorgestellt werden.

Konzeption und Umsetzung eines Themenmodells zur Generierung aktueller Themeninformationen

Grundlegende Herausforderungen für die Umsetzung künstlichen Themenbewusstseins in natürlichen Gesprächen ist die Identifikation der angesprochenen Gesprächsthemen und der den thematischen Ablauf eines Dialogs beeinflussenden Aktionen der Konversationsteilnehmer. Die dafür erforderlichen Prozesse werden innerhalb eines *Themenmodells* definiert, dessen Konzeption und technische Umsetzung im Rahmen der vorliegenden Arbeit durchgeführt werden. Resultierend aus der Motivation, das Themenmodell innerhalb der Mensch-Maschine-Interaktion einzusetzen und diese dadurch intuitiver zu gestalten, ergeben sich unterschiedliche Anforderungen, die es bei der Ausarbeitung des zugrunde liegenden Konzepts zu beachten gilt. Ausgehend von einem Anwendungsszenario, bei dem ein menschlicher und ein künstlicher Gesprächspartner (in Form eines konversationalen Agenten) in einer *face-to-face*-Situation miteinander kommunizieren, soll das resultierende Themenmodell folgende Anforderungen erfüllen:

- **Generalität** Das Themenmodell muss generell genug sein, um in nicht-aufgabenorientierten, domänenunabhängigen Dialogen eingesetzt werden zu können.
- **Echtzeitfähigkeit** Das Modell muss eine Verarbeitung neuer Gesprächsbeiträge zur Laufzeit der Interaktion gewährleisten, um das System jederzeit mit Informationen über die aktuelle Themensituation ausstatten zu können.

Die Anforderung, das Modell generell zu konzipieren, setzt den Zugang zu Wissen über Konzepte und mögliche Gesprächsthemen voraus. Die Online-Enzyklopädie *Wikipedia* stellt eine gewaltige Menge (aktuell über 1,4 Millionen in der deutschen Version¹) an bereits vorhandenen Beschreibungen unterschiedlichster Konzepte in Form von kollaborativ verfassten Artikeln bereit (Gabilovich & Markovitch, 2007). Die zusätzliche Einordnung dieser Artikel in eine ebenso kollektiv erstellte Kategorienstruktur reflektiert die menschliche Wahrnehmung von Entitäten und ihre thematischen Zusammenhänge. Eine adäquate Anbindung des Systems an diese Wissensquelle, deren

¹http://de.wikipedia.org/wiki/Deutschsprachige_Wikipedia, Stand: 19. Oktober 2012

Pflege und Instandhaltung in den Händen zahlreicher freiwilliger Autoren liegt, liefert demnach die Grundlage für die Erfüllung der Forderung nach Generalität. So können zuvor unbekannte Themen identifiziert werden, indem neue Gesprächsbeiträge auf bestehende Wikipedia-Inhalte abgebildet und in gemeinsamen Wikipedia-Kategorien zusammengefasst werden. Darauf aufbauend lässt sich eine *dynamische* Themenerkennung realisieren, welche unabhängig von der jeweiligen Domäne Informationen über die in dem Dialog angesprochenen Themen ermittelt.

Die Forderung nach Echtzeitfähigkeit verlangt eine Aktualisierung der Themeninformationen mit jedem neuen Gesprächsbeitrag. Zu diesem Zweck wird eine *online-fähige* Themenerkennung benötigt, die eine Erkennung des Gesprächsthemas ohne Zeitverzögerung erlaubt. Voraussetzung dafür liefern geeignete Vorverarbeitungsschritte, die das Wikipedia-Wissen zum einen auf einen Datensatz reduzieren, welcher nur die für die Realisierung des Themenmodells grundlegenden Informationen enthält, und zum anderen in einem Format ablegen, welches einen schnellen und kontinuierlichen Zugriff auf dieses frei verfügbare Wissen garantiert.

Um darüber hinaus den natürlichen Charakter kohärenter Konversationen aufzugreifen, werden die einzelnen Prozesse des Modells auf Grundlage fundierter und bewährter Theorien aus dem Bereich der Linguistik modelliert. Das Ergebnis ist eine *globale* Themenerkennung, die erst dann ein Gesprächsthema identifiziert, wenn es sich über mindestens zwei Dialogbeiträge erstreckt und somit die Betrachtungsweise eines Themas als ein von beiden Dialogpartnern gemeinsam durchgeführtes Projekt umsetzt.

Integration der gewonnenen Themeninformationen in das Dialogsystem des Agenten Max

Die aus der Themenerkennung resultierenden Themeninformationen gilt es anschließend auf den konversationalen Agenten *Max* (Kopp, Gesellensetter, Krämer & Wachsmuth, 2005) zu übertragen, um dessen bestehende konversationale Fähigkeiten um menschenähnliches Themenbewusstsein zu erweitern. Bei der Erweiterung des Dialogverhaltens um thematische Informationen sind folgende Anforderungen zu beachten:

- **Echtzeitfähigkeit** Die gegenwärtige Echtzeitfähigkeit des Agenten darf durch das zusätzliche Themenbewusstsein nicht beeinträchtigt werden.

- **Natürlichkeit** Das Wissen über Gesprächsthemen muss derartig in das Konversationsverhalten des Agenten einfließen, dass die Ausführung adäquater Äußerungen hinsichtlich der aktuellen Themensituation während der Interaktion mit Menschen gelingt.

Für eine Integration der Themeninformationen muss zum einen die Einbettung des Themenmodells in die Architektur des bestehenden Dialogsystems und zum anderen die Anpassung des konversationalen Verhaltens des Agenten hinsichtlich der aktuellen Themeninformationen gelingen. Bei der Einbettung des Themenmodells gilt es die ursprünglich auf Regeln basierende Architektur des Dialogsystems und somit die gegenwärtige Robustheit und Echtzeitfähigkeit des Systems zu bewahren. Die anschließende Anpassung des Konversationsverhaltens gelingt darauf aufbauend durch eine Erweiterung der bisherigen Dialogregeln von Max. Als Folge enthalten die Äußerungen des Agenten beispielsweise thematische Überschneidungen zu den vorherigen Gesprächsbeiträgen oder führen bei abrupten Wechseln des Gesprächsthemas zu einer entsprechenden Reaktion des Agenten, zum Beispiel zu einer Äußerung wie „*Wie kommst du denn nun auf das Thema?*“. Eine derartige Erweiterung der Dialogregeln simuliert den intuitiven Einfluss von Gesprächsthemen auf die Äußerungsproduktion und erfüllt die Forderung nach Natürlichkeit hinsichtlich des Gesprächsverhaltens des Agenten.

1.3 Das KnowCIT-Projekt

Die vorliegende Arbeit ist Bestandteil eines von Prof. Dr. Alexander Mehler aus dem Fachbereich der Linguistik und Prof. Dr. Ipke Wachsmuth aus dem Bereich der Informatik initiierten Projekts. Dieses ist dem Bielefelder Exzellenzclusters *CITEC* zugeordnet und trägt den Namen *KnowCIT* (Knowledge Enhanced Embodied Cognitive Interaction Technologies). Der Fokus des interdisziplinären Projekts liegt in der Anbindung des konversationalen Agenten Max an jüngste Webtechnologien, um die konversationalen Fähigkeiten des Agenten durch die Integration der online verfügbaren Informationen zu optimieren. Hierbei werden bestehende, kollaborativ erstellte Wissensbasen mit linguistischen *Information Retrieval*-Methoden ausgeschöpft und in die Architektur des Agenten integriert. Auf diese Weise erhält der Agent Zugriff auf Wissen, welches zum besseren Verständnis und zur thematischen Einordnung der Ge-

sprächsinhalte beiträgt und dazu genutzt werden kann, um Fragen des menschlichen Gesprächspartners auf Grundlage aktueller Informationen zu beantworten.

1.4 Aufbau der Arbeit

Für die Entwickler von Dialogsystemen ist es wichtig, über ein genaues Verständnis der Eigenschaften menschlicher Dialoge, wie sie strukturiert sind und wie Menschen sich an ihnen beteiligen, zu verfügen. Dies trifft insbesondere dann zu, wenn die Modellierung einer natürlich erscheinenden Konversation angestrebt wird (McTear, 2004). McTears Forderung nachkommend, vermittelt das nachfolgende Kapitel die für dieses Verständnis erforderlichen Grundlagen des zwischenmenschlichen Dialogs und beschäftigt sich mit dem Aufbau und der Fortführung sprachlicher Inhalte, den *Themen*. Danach liefert Kapitel 3 einen Überblick, wie sich die in Kapitel 2 diskutierten theoretischen Grundlagen auf künstliche Systeme übertragen lassen. So führt dieses Kapitel die Grundlagen automatischer Themenerkennung ein und präsentiert den aktuellen Stand der Forschung hinsichtlich künstlichen Themenbewusstseins in der Mensch-Maschine-Interaktion.

Nachdem in Kapitel 2 und 3 sowohl die theoretischen als auch die technischen Grundlagen vermittelt wurden, bilden die darauffolgenden Kapitel den Hauptteil und Schwerpunkt dieser Ausarbeitung. Gegenstand ist die Vorstellung des konkreten Konzepts und der genauen Umsetzung künstlichen Themenbewusstseins in natürlichen Dialogen zwischen einem virtuellen und einem menschlichen Gesprächspartner. Kapitel 4 führt diesbezüglich das Gesamtkonzept und die darauf aufbauende Architektur ein und stellt deren Einteilung in zwei Aufgabenschwerpunkte vor. Des Weiteren erfolgt für jeden der beiden Aufgabenschwerpunkte eine erneute, ausführlichere Darstellung der bereits in Abschnitt 1.2 zusammengetragenen Anforderungen und eine Diskussion der daraus resultierenden Konzeptentscheidungen.

Kapitel 5 beschäftigt sich anschließend mit der technischen Umsetzung eines Themenmodells zur automatischen Erkennung von Themen und thematischen Informationen in laufenden Dialogen. Dabei wird insbesondere der diese Erkennung durchführende Prozess vorgestellt, welcher die unterschiedlichen, den thematischen Verlauf eines Gesprächs beeinflussenden Aktionen identifiziert und die Resultate maschinenverständlich darstellt. Die abschließende Evaluation verdeutlicht das Potenzial der

automatischen Themenerkennung hinsichtlich ihres echtzeitfähigen Einsatzes in natürlichen Dialogen.

Kapitel 6 beschreibt die anschließende Integration der resultierenden Themeninformationen in das konversationale Verhalten des Agenten Max. Nach der Veranschaulichung der technischen Einbettung des Themenmodells in die bereits bestehende Systemarchitektur, wird die Anpassung des Dialogwissens beschrieben. Diese ermöglicht die Ausführung thematisch bedingter Äußerungen des Agenten. Abschließend wird ein Beispielszenario präsentiert, welches den durch das künstliche Themenbewusstsein des Agenten Max beeinflussten Dialogverlauf zwischen dem virtuellen und dem menschlichen Interaktionspartner aufzeigt.

Im Resümee werden die wichtigsten Aspekte und die Ergebnisse der gesamten Arbeit noch einmal zusammengefasst und diskutiert. Abschließend werden Ideen für weiterführende Arbeiten vorgestellt.

Teile dieser Arbeit wurden bereits vorab veröffentlicht in (Breuing & Wachsmuth, 2010b, 2010a; Breuing, Waltinger & Wachsmuth, 2011; Breuing & Wachsmuth, 2012).

2 Themen in Dialogen

Für die Emulation menschlichen Themenbewusstseins innerhalb eines interaktiven Systems ist ein genaues Verständnis der Beschaffenheit verbaler Kommunikation und insbesondere der menschlichen Fähigkeit, gemeinsam ein thematisches Gespräch aufzubauen und fortzuführen, erforderlich. Dieses Kapitel legt den Schwerpunkt auf die Vermittlung der entsprechenden Grundlagen und liefert die dieser Arbeit zugrunde liegenden Definitionen von elementaren Grundbegriffen. Nach einer allgemeinen Einführung in die Eigenschaften zwischenmenschlicher Kommunikation und der Darstellung des Dialogs als spezifische Kommunikationsform der gesprochenen Sprache in Abschnitt 2.1, beschäftigt sich Abschnitt 2.2 mit Themen als inhaltliche Gegenstände einer Kommunikation. Auf Basis bestehender Theorien wird die Bedeutung menschlichen Themenbewusstseins anhand der Darstellung des grundsätzlichen thematischen Gesprächsablaufs vermittelt. Anschließend werden in Abschnitt 2.3 die Strukturen der natürlichen Darstellung des Wissens über Konzepte und ihre Zusammenhänge in Form des mentalen Lexikons dargestellt. Das Kapitel schließt mit einer Zusammenfassung der zentralen Ausgangspunkte.

2.1 Sprachbasierte Kommunikation

Die Verwendung von Sprache als Methode zwischenmenschlicher Kommunikation gilt als eine der komplexesten Fähigkeiten der Menschen (Levelt, 1989). Schließlich hängt das Gelingen eines Informationsaustausches auf Sprachebene nicht nur von den individuellen Sprachhandlungen, sondern auch von der erfolgreichen Zusammenarbeit der teilnehmenden Kommunikationspartner ab. So stehen sowohl Sprecher als auch Empfänger im Zentrum jeder sprachlichen Handlung, deren einzelne Aktionen nicht nur auf den jeweiligen Gesprächspartner zugeschnitten, sondern auch aufeinander abgestimmt werden müssen (Clark, 1996). Sprachbasierte Kommunikation beruht somit auf individuellen, kognitiven, aber auch auf sozialen, koordinierenden Prozessen. Clark

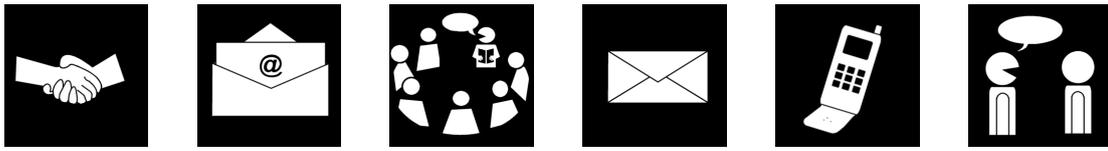


Abbildung 2.1: Verschiedene Kommunikationsarten auf Basis gesprochener und geschriebener Sprache (MDKMedia, 2012). Die beiden letzten Bilder repräsentieren typische Gesprächsformen.

(1996) deklariert den Gebrauch von Sprache daher als **joint action**, die auf einzelnen, individuellen Aktionen aufbaut.

Hinsichtlich der gegebenen Situation und des verwendeten Mediums lassen sich unterschiedliche Verwendungen von Sprache zur Kommunikation identifizieren. Beispiele hierfür sind Lektüren, Interviews, Aufsätze oder Vorträge (siehe Abbildung 2.1). Diese Formen des Sprachgebrauchs sind spezifiziert durch zusammenhängende Abfolgen situationsbezogener, sprachlicher Aktionen, wobei eine bestimmte Abfolge allgemein als **Diskurs** bezeichnet wird (Linell, 1998; Jurafsky & Martin, 2009). Diskurse lassen sich gemäß des zur Verständigung benutzten Kommunikationsmittels in *gesprochene* und *geschriebene* Diskurse unterteilen. Dabei stellt laut Levelt (1989) das **Gespräch** (engl. *conversation*), definiert als freie Interaktion zwischen zwei oder mehr Personen, die wohl ursprünglichste und universellste Form eines gesprochenen Diskurses dar: Ursprünglich, da die Wiege allen Sprachgebrauchs in dem Gespräch zwischen Kind und Eltern liegt, und universell, weil es den gesellschaftlich gängigsten Rahmen für die Verwendung von Sprache liefert (Levelt, 1989). So kommunizieren wir alltäglich auf Basis von Gesprächen, unter anderem, wenn wir im Restaurant unser Essen bestellen, mit unseren Freunden telefonieren oder uns über das Wetter beschweren (Jurafsky & Martin, 2009).

2.1.1 Dialog

In der Literatur wird der Begriff „Gespräch“ oft synonym mit dem Begriff „Dialog“ verwendet. Innerhalb der vorliegenden Arbeit wird der **Dialog** jedoch als Unterform des Gesprächs verstanden und ist definiert als kommunikative Handlung „zwischen mindestens zwei Personen, die simultan (am gleichen Ort oder verbunden durch einen technischen Kanal) miteinander so in Verbindung stehen, daß sie wechselseitig die Sprecherrolle bzw. die Hörerrolle übernehmen, wobei die Aufmerksamkeit auf gemein-

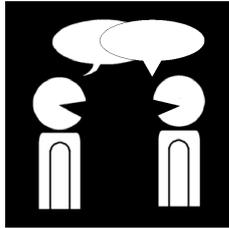


Abbildung 2.2: Der Dialog als wechselseitige Gesprächsform zwischen zwei Personen.

same, möglicherweise wechselnde Themen gerichtet ist.“ (Rath, 1979, S. 36,37). Ein Gespräch kommt dagegen ohne wechselseitige Interaktionen aus und kann beispielsweise in Form eines *Monologs* stattfinden. Abbildung 2.2 stellt eine Dialogsituation zwischen zwei Personen dar, während das Bild ganz rechts in Abbildung 2.1 als Darstellung eines Monologs interpretiert werden kann. Im weiteren Verlauf dieser Ausarbeitung bildet der Dialog die Grundlage für alle nachfolgenden konzeptionellen Überlegungen, die theoretischen Grundlagen jedoch beziehen sich weitestgehend auf die allgemeinere Kommunikationsform, das Gespräch.

Ein Dialog ist also eine auf Sprache basierende, zwischenmenschliche Interaktion, an der sich zwei oder mehr Personen mittels thematischen, meist gesprochenen Beiträgen beteiligen. Voraussetzung für das Gelingen eines Dialogs ist die Kooperation aller Beteiligten. Damit ist nicht zwangsläufig gemeint, dass die einzelnen Dialogpartner einer Meinung sein müssen, denn trotz großer Diskrepanzen kann ein Dialog gelingen. Vielmehr kommt es auf eine reibungslose Aushandlung der Verteilung der Dialogbeiträge zwischen den Teilnehmenden (siehe *turn-taking* (Sacks, Schegloff & Jefferson, 1974)) und auf Annahmen über ein gemeinsames Verständnis des Gesagten (siehe *common ground* (Clark & Schaefer, 1989)) an (McTear, 2004). Beide Phänomene werden unterstützt durch die von Rath (1979) in seiner Definition geforderte Aufmerksamkeit der Dialogpartner auf gemeinsame Dialogthemen, da Mehrdeutigkeiten und Referenzen aufgrund thematischer Eingrenzung leichter aufgelöst werden können und die Auswahl geeigneter Äußerungen leichter gelingt. Auch hierbei ist das Miteinander der Beteiligten gefordert, denn die Wahl eines Themas hängt, genau wie der gesamte Verlauf des Diskurses, von der Zusammenarbeit aller Dialogteilnehmer ab.

Menschen haben für gewöhnlich keine Schwierigkeiten, spontan einen Dialog mit anderen Personen aufzubauen, da sie darin eine gewisse Routine besitzen. Diese Routine wird während der Prozesse des Spracherwerbs und der Sozialisierung entwickelt

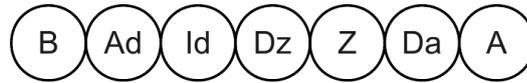


Abbildung 2.3: Die Elemente eines vollständig entwickelten Dialogs in Form einer zusammenhängenden Kette, in Anlehnung an Ventola (1979). Die Erklärung der einzelnen Elemente ist dem Text zu entnehmen.

und ist geprägt durch die jeweilige Kultur (Ventola, 1979). Erworben werden dabei Kenntnisse über Verhaltensmuster und Struktur von Dialogen, die es Personen erlauben, sich in unterschiedlichen Situationen angemessen zu verhalten. Diesbezüglich unterteilt Ventola einen Dialog in separate Elemente, die einer bestimmten Anordnung folgen (s. Abbildung 2.3). Demnach beginnt ein verbaler Diskurs in der Regel mit einer *Begrüßung* (*B*) (z. B. „Hallo“ oder „Guten Tag“), welche den Dialog eröffnet. Begrüßungen können verbal oder nonverbal erfolgen und verlangen zumindest in unserer Kultur eine entsprechende Rückantwort, ohne die ein reibungsloser Übergang zum nächsten Element erschwert wird. Eine Begrüßung wird häufig durch die Angabe der *Adresse* (*Ad*), das heißt des Empfängernamens, vervollständigt. Je nach Grad der Bekanntschaft handelt es sich dabei um den Vornamen (z. B. „Hallo *Christian*“) oder den Nachnamen (z. B. „Guten Abend *Herr Schmidt*“), der gegebenenfalls durch weitere Angaben, wie zum Beispiel einem akademischen Titel oder einer Berufsbezeichnung, ergänzt wird. Sind sich die Gesprächsteilnehmer nicht bekannt, so folgt zusätzlich eine Phase der *Identifikation* (*Id*), bei der sich die Sprecher gegenseitig vorstellen.

Ein weiteres Element in der Struktur kommunikativer Dialoge ist der *Dialogzugang* (*DZ*), der den eigentlichen Dialog durch sogenannte **safe topics** (Themen, die nicht zu persönlich und unangenehm sind, s. Kapitel 2.2.2) initiiert und für eine behagliche Gesprächssituation für alle teilnehmenden Personen sorgt. Nach Abschluss der bisher genannten, initialen Dialogelemente stehen zwei Optionen des weiteren Vorgehens zur Verfügung. Entweder beenden die Gesprächspartner ihren Diskurs und beschränken ihren Dialog somit auf einen *Smalltalk* beziehungsweise eine *lockere Unterhaltung* (engl. *casual conversation* (Ventola, 1979)), die nicht über die Thematisierung harmloser, meist oberflächlicher Inhalte hinausgehen, oder sie fahren mit einem tiefgründigeren Dialog fort. Im letzteren Fall beginnt nun die Phase der *Zentralisierung* (*Z*), bei der mindestens ein inhaltsreiches Thema im Mittelpunkt des Dialogs steht. Die Beendigung eines Diskurses wird eingeleitet durch einen *Dialogabschluss* (*Da*) in Form einer Äußerung, die in der Regel die Gründe für das Vorhaben, den laufenden

Dialog zu schließen, nennt. Im Anschluss erfolgt der *Abschied* (A), zum Beispiel durch ein „*Tschüss*“ oder „*Bis bald*“, der, wie die Begrüßung, um weitere Informationen ergänzt werden kann.

2.2 Themen in sprachlichen Diskursen

Einem Gespräch liegen typischerweise Gegenstände oder Inhalte zugrunde, sogenannte Themen, über welche sich die teilnehmenden Personen unterhalten. Nach Heritage und Watson (1979) bilden sie sogar die Grundlage für strukturierte Gespräche, da einzelne Diskurse auf Themen ausgerichtet organisiert sind. Die Anzahl existierender Arbeiten, deren jeweiliger Fokus auf die Erforschung von Themen in natürlichen Gesprächen liegt, ist zahlreich. Im Folgenden soll ein kurzer Abriss der wichtigsten Publikationen gegeben werden.

2.2.1 Der Begriff „Thema“

Der Begriff **Thema** leitet sich aus dem Griechischen und Lateinischen ab, wo er *Satz, abzuhandelnder Gegenstand* bedeutet. Ein Thema reflektiert gemeinhin, *worüber* sich Gesprächsteilnehmer unterhalten. Dabei sind Menschen dazu in der Lage, selbst längere Gesprächssequenzen in wenigen Worten oder sogar mit nur einem Begriff zusammenzufassen und als Thema zu benennen. Nach wie vor beschäftigt dieses Phänomen Wissenschaftler aus den verschiedensten Forschungsbereichen, die Anfänge der Erforschung von Gesprächsthemen erfolgten dabei aus soziologischer Perspektive. So veröffentlichten Landis und Burttt im Jahr 1924 eine Studie, in der die von einem gewissen Henry T. Moore zuvor aufgestellte These, die Wahl des Gesprächsthemas sei abhängig von dem Geschlecht der an dem jeweiligen Diskurs beteiligten Personen, untersucht wurde. Auf Grundlage von zehn Themenkategorien wie „Sport“, „Kleidung und Schmuck“, „Wetter“ oder „Gesundheit“ wurden unterschiedliche Gespräche zwischen Personen aus allen Alters- und Gesellschaftsgruppen in unterschiedlichen Situationen analysiert (Landis & Burttt, 1924). Dabei erfolgten sowohl die Festlegung der thematischen Kategorien als auch die Einordnung der einzelnen Gespräche in diese Kategorien durch die Autoren selbst.

Erst viele Jahre später traten Identifizierung und Klassifizierung von Themen in den Fokus der Linguisten und Kognitionswissenschaftler. Ziel der unterschiedlichen Ansät-

ze ist nach wie vor die Formulierung einer exakten Definition, welche die Form, grundlegenden Bestandteile und den Umfang eines Themas festlegt. Die erste Ausarbeitung, die eine solche Definition lieferte, erschien 1966 von Ellis (Ellis, 1966). Darin definiert der Autor ein Thema als *formale Kategorie* (engl. *formal category*) und beschreibt diese als Nominalphrase, die selbst oder aber referenziert in aufeinander folgenden Sätzen wiederkehrt und diese thematisch miteinander verbindet. Ellis' damalige Betrachtung eines Themas als dominierenden Gesprächsgegenstand eines Satzes, welcher sich über einen bestimmten Diskursabschnitt hinweg erstreckt, lässt sich noch heute in vielen grammatikalischen Ansätzen zur Themenbestimmung wiederfinden. Die analytischen Ansätze hingegen erlebten einen Wandel und betrachten nicht mehr einzelne Sätze, sondern komplette Diskurse, um auf Grundlage dieser das Gesprächsthema zu bestimmen (Downing, 2000). Dementsprechend unterscheidet van Dijk ein Satzthema (engl. *sentence topic*) von einem **Diskursthema**, wobei der von Keenan und Schieffeling (1976) eingeführte englische Begriff *discourse topic* ursprünglich das von Ellis vorgestellte Phänomen, dass sich ein Satzthema über mehrere Äußerungen hinweg erstreckt, bezeichnet. Van Dijks Ansatz, Sequenzen als Ganzes zu betrachten um ihre Themen zu bestimmen, wird von der Aussage Schanks, „*Sentences [...] do not have topics in isolation. Only conversations can be said to have topics*“ (Schank, 1977, S. 425), zusätzlich untermauert. Laut Schank stellen einzelne Sätze potentielle Themenvorschläge, sogenannte *topic candidates* (Linell, 1998), bereit, deren jeweilige Etablierung von der Akzeptanz des Gegenübers abhängt (Svennevig, 1999) (siehe *topic proffering*, Kapitel 2.2.2).

Eine weitere Unterscheidung lässt sich zwischen einem lokalem und einem globalen Thema treffen, um ein Thema bezüglich seines Fortbestands innerhalb des Gesprächs zu klassifizieren (Bublitz, 1989; Downing, 2000). So wird ein Thema als *lokal* bezeichnet, wenn es durch eine Äußerung vorgeschlagen wird und der andere Gesprächspartner dieses in seinem darauffolgenden Beitrag aufgreift (Bublitz, 1989). Ein *globales* Thema hingegen deckt den gesamten Diskurs oder zumindest eine Sequenz mehrerer Sprecherwechsel, sogenannte *Episoden* (Linell, 1998), ab. In der Regel existiert ein Thema also nicht schon vorab, sondern entwickelt sich durch eine wechselseitige Interaktion der am Gespräch Beteiligten (Downing, 2000). Clark (1996) bezeichnet ein Thema daher als **joint project**.

Doch was genau ist ein Thema, wie lässt es sich benennen? Ein Thema wird von

den Gesprächsteilnehmern unterschiedlich abstrakt und aus unterschiedlichen Perspektiven definiert (Ervin-Tripp, 1972; Svennevig, 1999). Die Autoren Brown und Yule (1983) verschafften der Spezifikation von Themen einen Spielraum, indem sie ein *topic framework* definierten, das den Bereich des geteilten, aktivierten Wissens der sich unterhaltenden Personen repräsentiert. *Speaking topically* wird somit verstanden als Formulierung von Äußerungen, deren angesprochenen Gegenstände innerhalb des aktiven Frameworks liegen (Brown & Yule, 1983). Schneider (1988) vertritt eine konkretere Auffassung und gibt zu bedenken, dass Begriffe, die als Bezeichnung für Themen gewählt werden, nicht unbedingt explizit innerhalb des Diskurses erwähnt werden, sondern für gewöhnlich *Supernyme*, also Überbegriffe, sind. Bublitz teilt diese Sichtweise: „*The topic is not identical with individual contributions to the discourse (or parts thereof), but it is a category distinguishable from them and superordinate to them*“ (Bublitz, 1989, S. 43).

Diskussion

Der Überblick existierender Theorien zeigt im Wesentlichen zwei unterschiedliche Perspektiven auf: die Betrachtung eines Themas als eigenständiges Element und diejenige als strukturierender und zusammenfassender Rahmen. Wird ein Thema als Element betrachtet, steht im Vordergrund, *was* besprochen wird (*aboutness* (Downing, 2000)), ein thematischer Rahmen impliziert hingegen, *wie* ein neues Thema angesprochen und weitergeführt wird (Goutsos, 1997). Bezüglich der ersten Perspektive wird ein Thema oft als Sachverhalt oder Kategorie definiert (Ellis, 1966; Keenan & Schieffelin, 1976; Bublitz, 1989). Letztere Perspektive konzentriert sich hingegen auf die allgemeine Diskursorganisation hinsichtlich des gewählten Themas. Sie geht mit der Auffassung einher, dass das, was gesagt wird, aus der Art, wie es gesagt wird, hervorgeht (Brown & Yule, 1983). Svennevig (1999) unterscheidet diesbezüglich die Betrachtung eines Themas als *Produkt* oder als *Prozess*, wobei der für die Differenzierung der beiden Betrachtungsweisen ausschlaggebende Faktor in dem Einfluss eines Themas auf das Zusammenspiel der beteiligten Personen liegt. So kann ein Thema als Produkt der wechselseitigen Zusammenarbeit der Gesprächspartner aufgefasst werden, oder aber als Menge an Techniken, die den Prozess der Zusammenarbeit unterstützen und den Diskurs zur Laufzeit organisieren (Svennevig, 1999).

Innerhalb der vorliegenden Arbeit wird ein Thema genau dann als Produkt be-

trachtet, wenn es im Zusammenhang mit der automatischen Erkennung von Themen einhergeht, da diese auf den vorliegenden Äußerungen des Dialogs operiert und auf Basis dieser Informationen über den temporären Gesprächsgegenstand ableitet. Bei der Emulation menschlichen Themenbewusstseins steht jedoch die Umsetzung kohärenter Dialoge zwischen dem künstlichen Agenten und einem menschlichen Gesprächspartner im Vordergrund, sodass für diese Herausforderung die übergeordnete Struktur des Dialogs wesentlich ist. Der Einbezug des thematischen Rahmens ermöglicht beispielsweise, den Agenten initiativ Themen vorschlagen lassen zu können. Der Agent wird somit zum aktiven Interaktionspartner und beeinflusst den Verlauf des weiteren Gesprächs genau wie sein menschliches Gegenüber.

Lag der Schwerpunkt zu Beginn der Erforschung von Themen auf der Analyse einzelner Sätze beziehungsweise Äußerungen, trat relativ rasch die Themenbestimmung für vollständige Diskurse in den Mittelpunkt. Bei der Untersuchung einzelner Sätze wird üblicherweise zwischen einem *Thema* und einem *Rhema* unterschieden. Ein Thema bezeichnet dabei das bereits Bekannte, welches typischerweise zu Beginn eines Satzes steht, im Gegensatz zum Rhema, welches die bisher unbekannt Information in der Mitte oder zum Ende eines Satzes darstellt. Die auf Sprache basierende Kommunikation scheint allerdings Einheiten zu umfassen, die länger sind als einzelne Äußerungen und obwohl Diskurse bezüglich ihrer Länge und Komplexität stark variieren, scheinen sie dennoch begrenzt und damit als natürliche Spracheinheit qualifiziert zu sein (Hurtig, 1977). Die Stärke der Analyse von Diskursen gegenüber Sätzen liegt in der Möglichkeit, Themen in Gesprächen zu verfolgen und sanfte Übergänge zu anderen Themen, sogenannte *topic drifts* (s. Kapitel 2.2.2) feststellen zu können. Im Rahmen dieser Arbeit wird die Erkennung von Themen daher auf Diskursebene realisiert. Ein Diskursthema spezifiziert dabei den Gesprächsgegenstand einer Sequenz an Dialogbeiträgen (Keenan & Schieffelin, 1976), sodass ein einzelner Diskurs mehrere Diskursthemen enthalten kann.

Eine weitere Kontroverse innerhalb der vorgestellten Theorien steht im Zusammenhang mit der Klassifizierung von Themen. Die von Ellis (1966) eingeführte Betrachtung eines Themas als angeführte Nominalphrase ist für die Klassifizierung eines Satzthemas hinreichend, impliziert jedoch nicht den übergeordneten Charakter eines Diskursthemas. Im Gegensatz dazu unterstreicht die Definition eines Themas als Kategorie eben diesen Charakter und erlaubt es, eine Sequenz an Diskursbeiträgen thematisch zusam-

menzufassen, indem es die besprochenen *Konzepte* subsumiert (Schank, 1977) und bis zu einem gewissen Grad generalisiert. Konzepte werden hier nach Gabrilovich und Markovitch (2007) verstanden, die sie als Basiseinheiten von Bedeutungen definieren, die den Menschen ermöglichen, ihr Wissen zu organisieren und zu teilen.

2.2.2 Themenbewusstsein

Thematische Dialoge sind das Ergebnis einer wechselseitigen Verhandlung der Beteiligten über das gemeinsame Dialogthema und basieren auf der Produktion und dem Austausch zusammenhängender Äußerungssequenzen. Grundlage für Letztere liefert das menschliche **Themenbewusstsein**, welches auf sowohl kognitiven als auch interaktiven Prozessen beruht. Es beschreibt die Fähigkeit, Themen sowohl wahrzunehmen als auch in das individuelle Gesprächsverhalten einzubeziehen. Dabei werden die individuellen Äußerungen durch vorherige Beiträge beeinflusst und beeinflussen wiederum nachfolgende Beiträge (Geluykens, 1999). Dies hat einen globalen Zusammenhang zwischen den einzelnen Äußerungen zur Folge, der gemeinhin als **Kohärenz** bezeichnet wird.

Kohärenz

In Anlehnung an die von Wolf und Gibson gelieferten, ursprünglich in englischer Sprache verfassten Beispieltex te zur Unterscheidung kohärenter und inkohärenter Satzsequenzen (Wolf & Gibson, 2006, S. 1), seien folgende Diskursausschnitte eines Dialogs zwischen Sprecher A und B gegeben:

- (1) A: „*Beim Raketenstart gestern war das Wetter gut.*“
B: „*Der Start der neuen Rakete konnte also wie geplant stattfinden?*“
A: „*Ja, sie bringt nun zwei Satelliten ins Weltall.*“

- (2) A: „*Ein neuer Satellit wurde ins Weltall geschossen.*“
B: „*Deshalb mag Sonja so gern Spinat!*“
A: „*John ist heute im Bett geblieben.*“

Dem Betrachter dieser Beispiele wird sofort ersichtlich, dass der erste Dialogabschnitt in sich stimmig und schlüssig ist, während der zweite Abschnitt eigenartig erscheint.

Warum ist das so? Die Äußerungen beider Beispiele sind grammatikalisch korrekt und leicht verständlich. Das Problem muss daher die Sequenz, also die Anordnung der Äußerungen betreffen (Jurafsky & Martin, 2009). Während die ersten beiden Beiträge des ersten Abschnitts beispielsweise in einem kausalen Zusammenhang stehen (das Wetter muss für einen Raketenstart generell gut sein; da das Wetter tatsächlich gut war, konnte der Raketenstart wie geplant durchgeführt werden), muss der Betrachter für den Aufbau eines kausalen Zusammenhangs im zweiten Beispiel eigenhändig einen Kontext zwischen den Äußerungen erdenken. Van Dijk spricht in diesem Fall von sogenannten *missing links*, die die Schwäche natürlicher Sprache, oft nicht voll explizit zu sein, hervorheben. Folglich erscheint der zweite Diskurs nur dann kohärent für den Betrachter, wenn dieser zusätzliche Informationen wie Hintergrundwissen oder das aktuelle Dialogthema beisteuert, die in den Äußerungen selbst nicht enthalten sind (Wolf & Gibson, 2006). Je leichter es für die jeweiligen Dialogteilnehmer ist, Zusammenhänge zwischen den einzelnen Äußerungen herzustellen, desto leichter verständlich ist der gesamte Diskurs. Demnach sind die Teilnehmer auf die Erkennung dieser Zusammenhänge und somit auf Kohärenz angewiesen, um einen Dialog zu verstehen und sich an diesem zu beteiligen (Jurafsky & Martin, 2009).

Während eines Dialogs einigen sich die am Diskurs beteiligten Personen interaktiv über den Inhalt ihres Gesprächs und erzeugen dabei Sequenzen kohärenter Äußerungen (Geluykens, 1999). Thematische Kohärenz wird daher, wie das Thema selbst, als gemeinschaftliche Errungenschaft aller Dialogteilnehmer verstanden (Downing, 2000; Geluykens, 1999). Dabei spielen nach Geluykens (1999) zwei Aspekte eine Rolle: a) Wird ein neues Thema eingeführt, muss der jeweilige Sprecher vorab einen möglichen Themenwechsel signalisieren. b) Sobald das neue Thema vorgeschlagen oder eingeführt wurde, muss es von den anderen Teilnehmern ausgehandelt und bestätigt werden, um in den laufenden Dialog integriert zu werden. Schlägt dieser Prozess fehl, wird keine thematische Kohärenz erreicht. Kohärenz im Dialog verzeichnet demnach einen höchst interaktiven Charakter und es wird pointiert verdeutlicht, „*that it indeed 'takes (at least) two to cohere' in terms of topic flow*“ (Geluykens, 1999, S. 37).

Zusammenhänge zwischen Sätzen, beziehungsweise zwischen Äußerungen, lassen sich nicht nur semantisch, sondern auch syntaktisch, durch sogenannte *Kohäsion*, herstellen. Kohäsive Hilfsmittel wie Wiederholungen von Schlüsselwörtern, die Verwendung von Pronomen, Synonymen und Relationen wie „sie“ und „er“ werden eingesetzt,

um Wörter und Satzglieder einer Äußerung mit Wörtern und Satzgliedern der anschließenden Äußerung zu verbinden (Jurafsky & Martin, 2009; Brown & Yule, 1983). Kohäsion kann als Bestandteil der Kohärenz betrachtet werden, wird im Rahmen dieser Arbeit jedoch separat aufgefasst und im Folgenden ausgeklammert.

In ihrer Ausarbeitung über die Vermeidung von Missverständnissen in alltäglichen Gesprächen mit Hilfe von Kohärenz betonen Bazzanella und Rossana (1997) die entscheidende Funktion von Kohärenz als Filter, welcher den Gesprächsteilnehmern hilft, eine Anzahl an Alternativen für die Interpretation von Äußerungen auszuschließen. Trotz dieser Einschränkung verbleiben mehrere Interpretationsmöglichkeiten, die sowohl passend als auch kohärent in der gegebenen Gesprächssituation sind. Es gilt nun, genau die Interpretation zu erkennen, die von dem Sprecher vorgesehen ist. Gelingt dies nicht und der Zuhörer denkt dennoch, die Äußerung richtig verstanden und interpretiert zu haben, produzieren die Beteiligten weiterhin kohärente Äußerungen, verfolgen dabei aber jeweils einen anderen Grundgedanken. Der Dialog wird so lange *parallel* fortgesetzt, bis keine weiteren Äußerungen, die kohärent zu der jeweiligen Interpretation sind, mehr gemacht werden können. In diesem Fall markiert ein Mangel an Kohärenz das Missverständnis und löst einen Reparaturprozess aus, das heißt die Wiederherstellung der Kohärenz als gemeinsame Grundlage für weitere thematische Beiträge. Die in diesem Zusammenhang durchgeführte Rekonstruktion dessen, was genau falsch gelaufen ist, ist meistens sehr einfach für die Beteiligten - anscheinend sind sie sich über mögliche Schwachstellen ihrer vergangenen Beiträge und Interpretationen bewusst.

Thematischer Ablauf eines Dialogs

Themenbewusstsein ermöglicht Menschen, neue Themen vorzuschlagen, laufende fortzuführen, thematische Wechsel zu erkennen, Themen beim Namen zu nennen und bereits abgeschlossene Themen wieder aufzugreifen. Mit diesen unterschiedlichen Aktionen nehmen sie Einfluss auf den Themenverlauf und haben damit unter anderem die Möglichkeit, ein Gespräch thematisch zu lenken. Innerhalb von Dialogen sind sie dabei jedoch von der Kooperation der anderen Dialogteilnehmer abhängig. Folglich lassen sich Themen in Dialogen nicht planen, da sie erst während eines Dialogs gemeinschaftlich erstellt werden (Downing, 2000). Die verschiedenen Aktionen, die den thematischen Ablauf eines Dialogs ausmachen, lassen sich nach Gardner (1987) in sechs

Kategorien einteilen. Nachfolgend werden diese zu vier allgemeineren Kategorien zusammengefasst und um eine weitere Kategorie ergänzt: die Benennung von Themen. Letztere hat keinen direkten Einfluss auf den Themenverlauf, stellt aber einen grundlegenden Prozess für das Bewusstsein von Themen dar, da eine Themenbezeichnung als Referenz fungiert, um beispielsweise auf den thematischen Inhalt vergangener Gespräche zu verweisen.

Themeneinführung Nach der für einen Dialog typischen Anfangsphase der Begrüßung, Vorstellung, etc. (s. allgemeiner Dialogaufbau, Kapitel 2.1.1) wird das erste Dialogthema eingeführt (Gardner, 1987). Genauer gesagt macht einer der teilnehmenden Sprecher einen Themenvorschlag, welcher anschließend von seinen jeweiligen Dialogpartnern akzeptiert werden muss, um sich als Dialogthema zu etablieren. Dieser Vorgang wiederholt sich beliebig oft innerhalb dieses Gesprächs, sodass ein Dialog meistens mehr als nur ein Thema behandelt.

Theoretisch gesehen ist die Liste möglicher Gesprächsthemen so lang wie der Katalog einer Bibliothek, in der Praxis beschränken sich Menschen jedoch in ihren Unterhaltungen auf eine bestimmte Auswahl an Themen (Kellermann & Palomares, 2004). Die Wahl eines Themas hängt dabei stark von der jeweiligen Gesprächssituation und den damit verbundenen Faktoren ab. Oft reicht es aus, nur einen einzelnen Faktor zu verändern, um einen komplett anderen Themenverlauf hervorzurufen. So unterhalten sich die gleichen Personen beispielsweise über andere Inhalte, wenn sie in beruflicher Umgebung kommunizieren, als wenn sie sich privat treffen. Ebenso lässt sich ein Gespräch bei einem Geschäftsessen in einem Restaurant thematisch stark abgrenzen von einer Unterhaltung zwischen Freunden, die sich ebenfalls in diesem Restaurant treffen. Laut Kellermann und Palomares (2004) hängt die Wahl des Themas primär von der persönlichen Beziehung zwischen den Gesprächsteilnehmern ab.

Maynard und Zimmermann (1984) untersuchten die Unterschiede in der Themenwahl zwischen Fremden und Personen, die sich vorab bereits kannten. Sie fanden heraus, dass Unbekannte ihre potentiellen Gesprächsthemen meistens aus der anfänglichen Kennenlernphase ableiten, während Bekannte eher Neuigkeiten austauschen und sich dabei auf die Historie gemeinsamer Interaktionen stützen. In seiner Untersuchung hinsichtlich der Themenwahl zu Beginn eines Gesprächs legte Schneider (1988) den Fokus auf *Smalltalks*, also auf spontane Alltagsgespräche, zwischen sich unbekanntem Perso-

nen. Anhand seiner Ergebnisse spezifiziert er drei verschiedene Situationskategorien, die als mögliche Smalltalk-Themen fungieren:

Die **unmittelbare Situation** (*immediate situation*) beinhaltet Themen, die Elementen des sogenannten *Frames* der gegebenen Situation entsprechen. Ein Frame wird hier nach Minsky (1975) definiert, das heißt als Datenstruktur für die Repräsentation stereotyper Situationen. Mögliche Themen eines Party-Frames wären beispielsweise Getränke, Musik, Gäste oder Essen.

Die **externe Situation** (*external situation*) beinhaltet Themen aus dem erweiterten Kontext der unmittelbaren Situation. Diese Kategorie ist die am wenigsten limitierte und besteht aus Themen wie Politik, Sport, Filme oder Nachrichten.

Die **Kommunikationssituation** (*communication situation*) beinhaltet persönliche Themen der Gesprächsteilnehmer, also deren Hobbies, Familien oder Berufe.

Laut Schneider (1988) bildet immer die unmittelbare Situation den Ausgangspunkt für ein erstes Smalltalk-Thema. Je nach Kultur und Situation wird das Gespräch anschließend mit Themen der externen oder der Kommunikationssituation fortgesetzt. Um den Einfluss der Kulturen der Gesprächspartner auf die Themenwahl detaillierter zu untersuchen, verglichen Endrass, Rehm und André (2011) Kennenlerngespräche zwischen Deutschen und Japanern. Sie fanden heraus, dass Japaner häufiger die unmittelbare Situation thematisieren als Deutsche, letztere dafür umso mehr über Themen der externen Situation sprechen. Ebenso präferieren die Deutschen private Themen, die die Kommunikationssituation betreffen, gegenüber Gesprächen über die unmittelbare Situation (Endrass et al., 2011), wobei der Anteil persönlicher und unpersönlicher Themen in einem Gespräch prinzipiell von dem Persönlichkeitsverhältnis zwischen den Gesprächsteilnehmern abhängt (Kellermann & Palomares, 2004). So sind beispielsweise bei einem ersten Zusammentreffen allzu persönliche Themen wie schwere Krankheiten, Sex oder Geldprobleme tabu (Schneider, 1988).

Bei der Wahl eines Themas sind jedoch nicht nur die zwischenmenschlichen Beziehungen zu den jeweiligen Gesprächspartnern von Bedeutung, sondern auch ihre Interessen. Folglich sprechen Personen meistens genau die Themen an, von denen sie meinen, dass sie interessant für die anderen Teilnehmer sein könnten (Chafe, 1994). Interessant sind Themen genau dann, wenn sie erstens neue Informationen bereithalten und zweitens von bestimmtem Wert für die Gesprächspartner sind. Was genau

Interesse auslöst, kann durch sogenannte *Topicalizer* getestet werden, die meistens in Form von Fragen einen Gegenstand ansprechen und dem jeweiligen Gegenüber als Gesprächsthema anbieten (Svennevig, 1999).

Themenfortführung Wurde ein Thema angesprochen, beziehungsweise vorgeschlagen, entscheidet die von dem nachfolgenden Sprecher gewählte Reaktion, ob dieses Thema fortgeführt wird und sich somit etabliert. Zu diesem Zweck steht dem Sprecher eine große Menge an unterschiedlichen Antwortmöglichkeiten zur Verfügung. Die Wahl einer Antwort kann jedoch nicht willkürlich getroffen werden und insbesondere dann, wenn der Sprecher beabsichtigt, das aktuelle Thema fortzuführen, ist die Auswahl deutlich begrenzt (Downing, 2000; Schank, 1977). In diesem Fall gilt es, die einen thematischen Dialog kennzeichnende *lokale Verbundenheit* zwischen den einzelnen Äußerungen herzustellen, die dadurch zustande kommt, dass die Dialogteilnehmer ihre Beiträge an die unmittelbar vorherige Äußerung anpassen (1999). Doch wie kann dies gelingen? Ist es ausreichend, wenn die nachfolgende Äußerung Aspekte des direkt zuvor Gesagten aufgreift? Tracy (1984) weist auf zwei Ansätze des Aufbaus einer lokalen Verbindung hin: Entweder, der aktuelle Sprecher geht auf den letzten Teil der vorherigen Äußerung seines Gesprächspartners ein (*lokaler Ansatz* nach Halliday und Hasan (1976)) oder er übernimmt den Leitgedanken des vorher Gesagten und baut diesen weiter aus (*globaler Ansatz* nach Kintsch und van Dijk (1978)). Tracy konnte nachweisen, dass insbesondere die Übernahme des Leitgedanken zum Einsatz kommt, um ein Thema fortzuführen und einen thematischen Dialog zu realisieren.

Van Dijk (1977) näherte sich diesem Thema aus der Perspektive der Diskursanalyse und beschäftigte sich mit der Frage, wie sich die lokalen Verbindungen zwischen aufeinander folgenden Sätzen, beziehungsweise Äußerungen, ermitteln lassen. Er stellte diesbezüglich vier mögliche Methoden vor: Aufbau einer kohäsiven Verbindung, Ergänzung der Äußerungen durch die Zunahme fehlender Verbindungen, Abbildung auf Frame-Wissen und der Nachweis globaler Kohärenz des Diskurses. Bei genauerer Betrachtung dieser Methoden lassen sich einige Schwachstellen ausmachen. Eine kohärente Verbindung zum Beispiel kennzeichnet nicht zwangsläufig die Aufrechterhaltung eines Themas, wie der folgende (im Original englische) Ausschnitt eines Beispieldialogs nach Richard (1980, S. 415) aufzeigt (Gardner, 1987):

A: „Wie viel hast du für die Bluse bezahlt?“

B: „Magst du sie? Ich habe sie bei Metro gefunden.“

Der Dialogwechsel ist kohärent und weist Kohäsion in Form von zwei Referenten („sie“, damit wird auf „Bluse“ verwiesen) im zweiten Beitrag auf. Allerdings verschiebt sich das eigentliche Thema „Kosten der Bluse“ nach „Ort des Einkaufs“. Kohäsive Verbindungen zwischen Äußerungen sind demnach nicht nur bei Themenfortführungen gegeben, sondern liegen auch bei thematischen Verlagerungen (Gardner, 1987) vor (s. nächster Abschnitt). Demzufolge ist der Nachweis einer kohärenten Verbindung kein Garant für die Aufrechterhaltung des aktuellen Themas.

Lassen sich keine thematischen Übereinstimmungen zwischen zwei hintereinander auftretenden Beiträgen erschließen, bietet sich eine intuitive Ergänzung der enthaltenen Sachverhalte an, um mögliche Verbindungen zu vorherigen Äußerungen zu ermitteln. Eine anschließende Kategorisierung der dadurch sichtbar werdenden Themenüberschneidungen ermöglicht, Themenfortführungen von thematischen Wechseln und Wiederholungen zu unterscheiden. Gardner (1987) weist darauf hin, dass diese jedoch rein intuitiv erfolgen und von subjektiven Hypothesen hinsichtlich der Bedeutung der einzelnen Äußerungen abhängig ist. Subjektivität ist auch die Schwachstelle der nächsten Methode, die sich auf vorhandenes Frame-Wissen der Dialogteilnehmer stützt. Können die innerhalb des betrachteten Dialogwechsels enthaltenen Sachverhalte demselben Frame zugeordnet werden, besteht ein thematischer Zusammenhang zwischen den einzelnen Äußerungen. Die Probleme dieser Methode liegen laut Gardner in der Vielzahl an Möglichkeiten, Situationswissen in Frames abzulegen, die von den individuellen Wissensstrukturen der einzelnen Dialogteilnehmer abhängt. Ein ebenfalls rein intuitiver Ansatz ist die Identifikation globaler Kohärenz. Lässt sich ein den Dialogwechsel umgebender Diskursabschnitt vollständig einem gemeinsamen Thema zuordnen, existiert auch eine thematische Überschneidung zwischen den betrachteten Äußerungen.

Die lokale Verbundenheit von Dialogabschnitten schränkt sowohl die Produktion als auch die Interpretation anschließender Äußerungen ein. Dies gelingt, da jeder Teilnehmer zunächst von der Aufrechterhaltung eines Themas ausgeht. Störungen dieser müssen von dem jeweiligen Sprecher daher speziell angekündigt werden (Svennevig, 1999). In der Regel wird ein Thema so lange fortgesetzt, bis alle an dem Dialog Betei-

ligten nichts mehr zu diesem beizutragen haben.

Themenwechsel Während eines Dialogs werden für gewöhnlich mehrere unterschiedliche Gesprächsgegenstände thematisiert. Ein **Themenwechsel** erfolgt dann, wenn sich die Dialogpartner kollektiv einem neuen Thema zuwenden und das vorherige aufgeben (Svennevig, 1999). Es existieren zwei verschiedene Techniken, um ein Thema zu wechseln: Erstens, das aktuelle Thema wird zuerst von allen Teilnehmern gemeinsam abgeschlossen und anschließend ein Spielraum für die Einführung eines neuen Themas bereitgestellt. Zweitens, das neue Thema wird bereits vorgeschlagen, während das aktuelle Thema eigentlich noch in Bearbeitung ist (Svennevig, 1999). Auch wenn die erste Technik die effizientere und höflichere ist, da sie zum einen erkennbare Grenzen zwischen den einzelnen Themen erkennen lässt, die zum Verständnis neuer Äußerungen beitragen, und zum anderen allen Beteiligten die Möglichkeit gibt, ihre Beiträge zu einem Thema vollständig zu äußern, bevor ein neues Thema eingeleitet wird (Goodenough & Weiner, 1978), ist die zweite Technik diejenige, die am meisten Verwendung findet. Dementsprechend sind die Übergänge zwischen den einzelnen Themen in Gesprächen meist fließend und die Grenzen zu den neuen Themen nur schwer ermittelbar (Button & Casey, 1985).

Es existieren jedoch auch erkennbare Zeitpunkte innerhalb eines Gesprächs, an denen ein Themenwechsel eingeleitet wird und alle Beteiligten die Möglichkeit haben, das Thema kollektiv zu schließen, also ein Themenwechsel mit Hilfe der ersten Technik zu praktizieren, bevor sie mit einem neuen Thema fortfahren. Zum Beispiel, wenn in den Beiträgen der Gesprächspartner nichts mehr zu dem aktuellen Thema hinzugefügt wird und die Pausen zwischen den einzelnen Äußerungen länger werden. Dieser Mangel an *Progressivität* (engl. *progressivity* (Svennevig, 1999)), also an weiterer Ausführung und Entwicklung eines Themas, lässt sich insbesondere in gesprochenen Diskursen an formulierten Wiederholungen, Zusammenfassungen und Schlussfolgerungen der Sprecher erkennen und führt zu einer Beendigung des aktuellen Themas (Howe, 1991; Svennevig, 1999). Für gewöhnlich erstreckt sich diese Art des Themenwechsels über mehrere Äußerungen hinweg und wird somit kooperativ durchgeführt (Howe, 1991).

Demgegenüber stehen Themenwechsel, deren Übergänge durch kohärente Äußerungen und fließende Überleitungen zu neuen Themen gekennzeichnet sind. Gegeben sei folgende Beispielsequenz:

A: „In welcher Stadt lebst du?“

B: „München.“

A: „Ach, dann bist du bestimmt Fan von Bayern München.“

A: „Nein, eigentlich nicht. Ich mag Schalke 04.“

Mit seiner Frage schlägt Sprecher A vor, sich über Orte zu unterhalten. Sprecher B akzeptiert dieses Thema, indem er mit dem Namen seines Wohnortes antwortet. Obwohl dieser an sich keinen direkten Bezug zum Thema „Fußball“ erkennen lässt, ist die Stadt „München“ dennoch aus konzeptioneller Sicht eng verbunden mit dem örtlichen Fußballverein „Bayern München“. Dies hat zur Folge, dass der von Sprecher A initiierte Wechsel zum Thema „Fußball“ von Sprecher B akzeptiert und der Dialog nahtlos fortgesetzt wird, indem auch Sprecher B mit der nachfolgenden Äußerung auf das neue Thema eingeht. Diese Art des Themenwechsels wird nach Hobbs (1990) als **Themendrift** (engl. *topic drift*) und nach Hurtig (1977) als **Themenüberblendung** (engl. *topic fading*) bezeichnet und beschreibt das Phänomen, durch einen graduell verlaufenden Übergang einen innerhalb des laufenden Gesprächs mitunter radikalen Wechsel des Ursprungsthemas zu vollziehen. Hobbs vergleicht eine aus Themendriffs bestehende Konversation mit einem Wortspiel, bei dem ein vorgegebenes Wort in ein anderes vorbestimmtes Wort überführt werden soll, indem pro Runde ein einzelner Buchstabe geändert wird, sodass die Zwischenergebnisse wiederum Wörter sind. Dieser Prozess entspricht der Auffassung von Schank (1977), für den eine Äußerung sowohl Überlappungen zum aktuellen Thema, als auch Vorschläge für neue Themen enthalten muss. So kann ein Gespräch thematisch ganz woanders enden, als es angefangen hat, ohne dabei an einer Stelle die Kohärenz zu vernachlässigen. Dies gilt auch für die sogenannte **Themenschattierung** (engl. *topic shading* (Hurtig, 1977)), bei der das Thema nicht gewechselt, sondern expandiert wird. In Bezug auf den oben genannten Beispieldialog könnte eine weitere Äußerung von Sprecher A wie „Und, treibst du selbst auch Sport?“ eine Themenschattierung bewirken, da das aktuelle Thema „Fußball“ auf das allgemeinere Thema „Sport“ ausgedehnt werden würde.

Ein neues Thema kann aber auch abrupt, das heißt ohne konzeptionelle Verbindungen zu dem zuvor Gesagten und ohne Prozesse der Themenbeendigung, eingeleitet werden und zu einem *sprunghaften* Themenwechsel führen. In Anlehnung an Svenne-

vig (1999, S. 164) wird solch ein Themenwechsel als **Themensprung** (engl. *topic leap*) bezeichnet. Auch hier hängt der tatsächliche Wechsel und damit die Etablierung des neuen Themas von der Akzeptanz der jeweiligen Gesprächspartner ab. Allerdings muss ein Themensprung nicht zwangsläufig ein neues Thema einleiten. Da das alte Thema nicht kollektiv abgeschlossen wurde, bleibt es zunächst noch bestehen und abrufbar. Dies gibt den Gesprächsteilnehmern die Möglichkeit, auf dieses Thema zurückzukommen und den Wechsel somit als zeitliche Verschiebung des alten Themas aufzufassen. Sequenzen, die diese Art der zeitlichen Verschiebung eines laufenden Themas hervorrufen, sind Rand- oder Nebenbemerkungen und werden von Svennevig (1999, S. 186) als *off topic*, also als „thematisch unpassend“ bezeichnet. Beiträge, die *off topic* sind, weisen keine Kohärenz mit dem vorherigen Thema auf, passen nicht in die globale Organisation des Gesprächs und tragen nicht zum Fortgang des Dialogs bei.

Themensprünge können zum Beispiel durch Fragen eingeleitet werden. Fragen sind grundsätzlich als interaktiv zu bewerten, da sie in der Regel den ersten Teil eines Frage-Antwort-Paares darstellen (Sacks et al., 1974). Entsprechend hoch ist die Wahrscheinlichkeit, dass ein durch eine Frage eingeleiteter Themenwechsel gelingt. Geluykens unterscheidet zwei Arten von Fragen, die ein neues Thema einführen: **Themenvorschläge** (engl. *topic-proposing*) und **Themenentlockung** (engl. *topic-eliciting*) (Geluykens, 1999, S. 37). Bei einem Themenvorschlag stellt der Gesprächsteilnehmer, der ein neues Thema einleiten möchte, eine Frage zu diesem Thema und spricht es somit selbst an. Eine Themenentlockung hingegen „entlockt“ das gewünschte Thema dem nächsten Teilnehmer, indem der Sprecher eine relativ offene Frage stellt, welche explizit nach einem neuen Thema fragt. Wie bei jedem Themenvorschlag ist ein erfolgreicher Themenwechsel jedoch nicht garantiert, da dieser von der Kollaboration aller Gesprächsteilnehmer abhängig ist (Geluykens, 1999).

Themenwiedereinführung Weist eine Äußerung keine thematische Verbindung zu dem direkt vorherigen Beitrag auf, kann aber einem zuvor besprochenen, bereits abgeschlossenen Thema zugeordnet werden, so spricht man von einer potentiellen Themenwiedereinführung. Diese wird in der Regel speziell angekündigt, zum Beispiel durch eine Aussage wie „*Ich möchte nochmal auf das Thema XY zurückkommen.*“ (Gardner, 1987). Eine Wiederaufnahme eines Themas kann aber auch kommentarlos erfolgen, insbesondere, wenn ein thematischer Dialog unerwartet unterbrochen wurde, beispielsweise

se durch eine Nebenbemerkung (s. oben). Die Wiederaufnahme eines Themas hängt, wie jede Themenentstehung, von der Akzeptanz der restlichen Dialogpartner ab.

Themenbenennung Svennevig (1999) spezifiziert ein Dialogthema als individuell gewählte Kategorie. Die Bezeichnung eines Themas wird durch diese Wahl bedingt und lässt sich daher ebenfalls als Ergebnis individuell erfolgreicher Prozesse der Dialogteilnehmer zurückführen. Dabei fällt auf, dass nominale Formen für die Themenbenennung bevorzugt werden. Diese Feststellung stützt sich auf eine Untersuchung von Schneider (1988), der seine Studenten einen thematischen Diskurs analysieren ließ. Dabei wurden sieben der acht identifizierten Themen mit Nomen benannt. Abgesehen von der Präferenz für Nomen zeigte sich bei der Studie zudem eine große Übereinstimmung bezüglich der gewählten Bezeichnung, obwohl dieser nicht zwangsläufig explizit während des Diskurses erwähnt wurde. Schneider weist darauf hin, dass wenig über Themenbenennung bekannt ist, jedoch angenommen werden kann, dass die resultierenden nominalen Formen mit der *mentalen Repräsentation* der Dialogthemen (s. nachfolgender Abschnitt) korrelieren.

2.3 Das mentale Lexikon

Menschen übertragen die reale Welt in Konzepte, die es ihnen ermöglichen, ihr Wissen zu organisieren und gemeinsam zu benutzen. Denkt man beispielsweise an die Wörter „Tiger“ oder „Mond“ wird deutlich, wie genau die jeweils assoziierten Konzepte die externe Welt reflektieren, da es eine beachtenswerte Übereinstimmung hinsichtlich ihres Verständnisses gibt, sogar zwischen Menschen unterschiedlicher Sprachen (Aitchison, 2003). Basis für die Repräsentation von Konzepten in Form von Wörtern, ihren Bedeutungen, sowie zusätzlichen Informationen wie Aussprache und Zusammenhänge, liefert das innere Wörterbuch, **mentales Lexikon** genannt. Genau wie ein herkömmliches, geschriebenes Lexikon enthält es einen Index mit allen, dem jeweiligen Menschen bekannten Wörtern (schätzungsweise zwischen 150.000 und 250.000 (Aronoff, 2003)). Wann immer wir etwas in einer uns bekannten Sprache hören oder lesen, oder wir selbst Sprache produzieren, greifen wir auf diesen internen Bestand zu. Die Einzelheiten der Repräsentation und der Prozesse des Gebrauchs sind nicht vollständig verstanden, jedoch weist die erstaunliche Geschwindigkeit der Wortfindung und -erkennung (in-

nerhalb einer Fünftel Sekunde vom Beginn des Wortes bei gesprochener Sprache) auf ein höchst effizientes und organisiertes System hin (Aronoff, 2003).

Unterschiedliche Theorien hinsichtlich der Organisation des mentalen Lexikons aus den Bereichen der kognitiven Psychologie und der Künstlichen Intelligenz modellieren die enthaltenen Konzepte als Elemente, die in Form von assoziativen Verbindungen miteinander in Beziehung stehen (Schreuder & Flores d'Arcais, 1989). Ein sehr bekanntes, auf dieser Idee aufbauendes Modell, *Spreading Activation Network* genannt, wurde von Collins und Loftus (1975) entwickelt, welches die Bedeutung eines Wortes als Konzeptknoten mit Verbindungen zu anderen Knoten repräsentiert. Bei Zugriff auf einen Knoten dehnt sich die Aktivierung auf benachbarte Konzepte des Netzwerks aus, wobei näher gelegene, semantisch ähnlichere Knoten stärker aktiviert werden als entferntere.

Die Darstellung des mentalen Lexikons als Netzwerkstruktur wird auch von Aitchison (2003) übernommen, wobei die einzelnen Konzeptknoten in semantischen Feldern organisiert, das heißt, thematisch gebündelt sind. Innerhalb dieser Felder weisen die Verbindungen eine hohe Stärke auf, insbesondere dann, wenn die betrachteten Wörter derselben Kategorie angehören und den gleichen Detaillierungsgrad besitzen (wie zum Beispiel „Salz“ und „Pfeffer“ oder „blau“ und „grün“). Diese Struktur hilft unter anderem, die Suche nach speziellen, unüblichen Wörtern semantisch einzuschränken. Bestätigt wird dies durch Beobachtungen Aitchisons (2003), bei denen Personen häufig den semantischen Bereich und die Kategoriengruppe des gesuchten Wortes abtasten, um sich beispielsweise an den Namen eines bestimmten Konzepts zu erinnern. Entsprechend riefen Versuchspersonen, die das Wort „Sextant“ vergessen hatten, andere Navigationsinstrumente auf, wie zum Beispiel „Kompass“, um sich an das vergessene Wort zu erinnern. Ähnliche Phänomene lassen sich bei Personen mit Sprachstörungen beobachten. Sie erzeugen oftmals andere Wörter als ursprünglich gemeint, diese gleichen sich aber meistens bezüglich des Detaillierungsgrads mit der Bedeutung des ursprünglich geplanten Wortes oder weisen zu diesem eine kategorische Verwandtschaft auf. Folglich sagen sie oft „Orange“ statt „Zitrone“ oder „Tisch“ statt „Stuhl“. Diese Art des falschen Gebrauchs lässt eine enge Verbundenheit einiger Wörter gleicher Kategorie vermuten, die sich dadurch nur schwer voneinander unterscheiden lassen. Daher fällt es Probanden in der Regel schwerer, ein Bild mit einer Zitrone aus einer Menge an Bildern mit Obstsorten herauszusuchen, als aus einer Menge an Bildern anderer

Kategorien. Durch Studien mit älteren Personen ließ sich nachweisen, dass gerade diese starken Verbindungen zwischen Wörtern gleicher Kategorie nicht verlernt werden und selbst dann erhalten bleiben, wenn die betroffene Person zuvor schwere Hirnschäden erlitt (Aitchison, 2003).

Außergewöhnliche Verhaltensweisen von apathischen Patienten unterstreichen die semantische Organisation von Konzepten nach Themen. Auf die Frage nach der Bedeutung des Begriffs „Nadel“ antwortete ein Patient, er habe es vergessen. Auch die Wörter „Möhre“ und „Moskito“ konnte der Patient nur unzulänglich beschreiben. Aitchison (2003) berichtet weiter, dass derselbe Patient jedoch keine Schwierigkeiten hatte, einen „Pakt“ als freundschaftliche Übereinkunft, ein „Bittgesuch“ als ernsthafte Bitte um Hilfe und „Wissen“ als mentale Bekanntmachung mit einem Gegenstand zu definieren. Ein weiterer Patient konnte keine Früchte benennen, kannte aber die Namen von sämtlichen Küchenutensilien. Beide Personen beherrschten demnach Begriffe bestimmter Themenfelder, anderer jedoch nicht. Für Aitchison (2003) belegen diese Fälle, dass thematische Bereiche teilweise unabhängig voneinander abgelegt werden und einige semantische Felder daher beschädigt werden können, ohne andere zu beeinträchtigen.

Über diese Art der Wissensorganisation hinaus, kommt das mentale Lexikon einer zusätzlichen Strukturierung nach, der sogenannten *Hyperonymie*. Sie beschreibt eine *is a*-Relation zwischen zwei Begriffen, wobei der jeweils allgemeinere, das **Hyperonym** (zum Beispiel „Obst“), einen Oberbegriff des spezielleren Begriffs, des **Hyponyms** (zum Beispiel „Apfel“), darstellt (Aronoff, 2003). Das Hyponym erbt dabei alle Eigenschaften des allgemeineren Konzepts und ergänzt mindestens eine Eigenschaft, welches es von seinem Oberbegriff und den weiteren Hyponymen des Oberbegriffs abgrenzt (Miller, Beckwith, Fellbaum, Gross & Miller, 1990). So fungiert der Begriff „Eule“ als Oberbegriff von „Schneeeule“ und „Schleiereule“, ist selbst aber wiederum das Hyponym von „Vogel“. Letzteres ist ebenfalls in einer höheren Kategorie enthalten: Zusammen mit Fischen, Insekten und so weiter formen Vögel die Kategorie „Tiere“. Hyperonymie-Relationen sind transitiv und asymmetrisch, das bedeutet, dass „Schneeeule“ ein Hyponym von „Tiere“ ist, „Tiere“ jedoch kein Hyponym von „Schneeeule“ ist. Die resultierende, hierarchische Struktur lässt sich als Baum visualisieren und wird gemeinhin als **Taxonomie** (s. Kapitel 3.1) bezeichnet. Das mentale Lexikon folgt dieser Taxonomiestruktur, wenn auch nicht in Form des dargestellten Baumes (Aitchison,

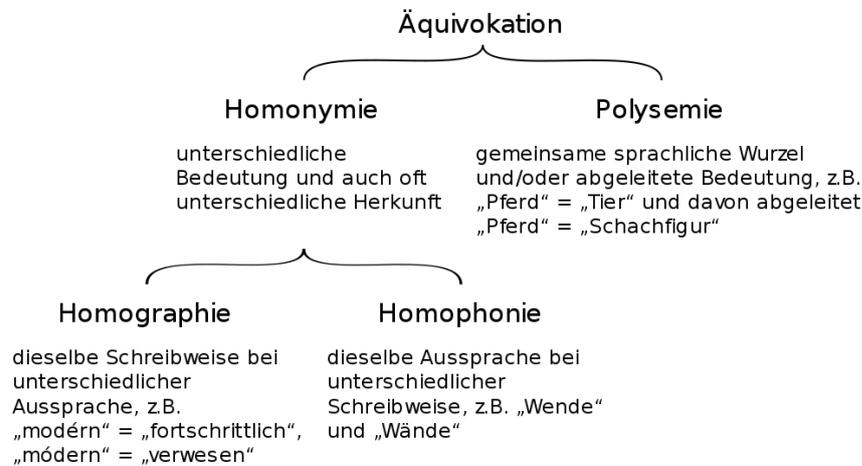


Abbildung 2.4: Lexikalische Mehrdeutigkeiten in der natürlichen Sprache. Quelle: Wikipedia (<http://de.wikipedia.org/wiki/Homonym>).

2003).

Wie bereits erwähnt werden Konzepte durch Wörter bezeichnet und repräsentiert. Unglücklicherweise existiert innerhalb der natürlichen Sprache keine bijektive Beziehung zwischen Wörtern und Konzepten, was zu **lexikalischen Mehrdeutigkeiten** (engl. *lexical ambiguities*) führt (Helbig, 2006). Diesbezüglich lassen sich aufgrund eines historischen Kriteriums zwei Phänomene unterscheiden, die **Homonymie** und die **Polysemie** (s. Abbildung 2.4). Polysemie bezeichnet das Vorkommen eines Wortes, dessen Bedeutung sich im Laufe der Zeit aufgespalten hat (Linke, Nussbaumer & Portmann, 2004). So lassen sich die beiden unterschiedlichen Bedeutungen „Schließvorrichtung“ und „Gebäude“ des Wortes „Schloss“ auf eine gemeinsame Grundbedeutung zurückführen. Im Gegensatz dazu beschreibt Homonymie die Mehrdeutigkeit von Wörtern, die ursprünglich ungleich lautenden *Morphemen*, das heißt den kleinsten, bedeutungstragenden Einheiten, entstammen und im Laufe der Geschichte äquivalent geworden sind. Linke, Nussbaumer und Protmann (2004) nennen das Wort „Kiefer“ als Beispiel für ein **Homonym**, welches zum einen aus dem mittelhochdeutschen Wort *kiver* hervorgegangen ist und heute die maskuline Form, also „der Kiefer“ darstellt, und zum anderen dem althochdeutschen Wort *kienforha* erwachsen ist und heute als feminine Form „die Kiefer“ bezeichnet. Wie aus der Abbildung 2.4 hervorgeht, lässt sich Homonymie überdies unterteilen in **Homographie** und **Homophonie**. Während *Homographen* in der Regel eine unterschiedliche Aussprache bei gleicher Schreibweise

(„die *Sucht*“, „er *sucht* seine Schlüssel“) aufweisen, klingen Homophone gleich, werden aber anders geschrieben („die *Lehre*“, „die *Leere*“).

Viele Wörter besitzen also mehrere Bedeutungen und weisen somit lexikalische Mehrdeutigkeit auf. Ende der 70er Jahre beschäftigte sich der Psycholinguist David Swinney mit der Frage, wie Menschen zwischen den unterschiedlichen Auswahlmöglichkeiten entscheiden und abhängig von dem jeweiligen Kontext die richtige Bedeutung des mehrdeutigen Wortes ermitteln. Unter Anwendung einer neu entwickelten Methode, der sogenannten *Cross-Modal Priming Task*, untersuchte Swinney (1979), ob bei der Verarbeitung mehrdeutiger Wörter zunächst alle etwaigen Bedeutungen aktiviert und durchsucht werden, um die eine richtige zu erhalten, oder ob nur die für den aktuellen Kontext geeigneten Bedeutungen als mögliche Kandidaten herangezogen und die Auswahlmöglichkeiten somit von vornherein eingegrenzt werden (Altmann, 1997). In der Studie wurden den vor einem Computermonitor sitzenden Probanden aufgezeichnete Sätze vorgespielt, die Wörter mit zwei unterschiedlichen, gleich häufig verbreiteten Bedeutungen enthielten. Sobald eines dieser mehrdeutigen Wörter genannt wurde, erschien eines seiner möglichen Bedeutungen auf dem Bildschirm (im Deutschen würde beispielsweise für das Wort „Schloss“ entweder „Gebäude“ oder „Schließvorrichtung“ auf dem Monitor erscheinen), ein Wort ohne jeglichen Bezug, oder aber ein nicht existierendes Wort. Die Aufgabe der Probanden war es dann, so schnell wie möglich zu entscheiden, ob es sich bei dem angezeigten Zeichenfolge um ein Wort handelt oder nicht. Neben den angezeigten Wörtern variierte die bereitgestellte Kontextinformation, sodass einige Sätze den jeweiligen Probanden stark in die Richtung einer bestimmten Bedeutung beeinflussten, andere Mehrdeutigkeiten jedoch nicht so eindeutig aufgelöst werden konnten. Aufgrund der stets gleich bleibenden Verarbeitungszeiten, die die Probanden für ihre Entscheidungen benötigten, konnte Swinney nachweisen, dass bei Mehrdeutigkeiten zunächst alle Alternativen möglicher Wortbedeutungen aktiviert werden, um Wörter als existent oder erfunden zu bestimmen, auch wenn das betrachtete Wort durch den Satz, in dem es eingebettet ist, eindeutig vorherbestimmt ist. Einige nachfolgende Studien konnten dieses Ergebnis nicht reproduzieren, wenn sie zwei zusätzliche Faktoren betrachteten: den unterschiedlichen Grad des Einflusses von Kontext und die unterschiedliche Verwendung der Bedeutungen (Altmann, 1997). Unterstützt der Kontext die weniger übliche Bedeutung, werden alle Bedeutungsmöglichkeiten aktiviert; wird jedoch die häufiger verwendete Bedeutung vom Kontext her

angesprochen, weist die weniger übliche Bedeutung wenig bis gar keine Aktivierung auf (Duffy, Morris & Rayner, 1988). In der Ausarbeitung von Altmann sind weitere Prozesse beschrieben, die zu jeder Zeit eine Reihe an Wortbedeutungen aktivieren (siehe (Altmann, 1997, Kapitel 6)), bis jetzt ist jedoch noch ungeklärt, warum wir Menschen uns nicht über diese Vielzahl der Bedeutungsaktivierungen bewusst sind.

2.4 Zusammenfassung

In diesem Kapitel wurden die für eine Emulation menschenähnlichen Themenbewusstseins in konversationalen Agenten erforderlichen theoretischen Grundlagen hinsichtlich der menschlichen Fähigkeit, sich eines Gesprächsthemas bewusst zu sein, vermittelt. Dabei wurden zunächst die Begriffe „Dialog“ und „Thema“ im Kontext zwischenmenschlicher Kommunikation eingeführt, um darauf aufbauend die einzelnen Aspekte menschlichen Bewusstseins für Themen in natürlichen Dialogen darstellen zu können. Dabei wurden insbesondere der thematische Ablauf eines Dialogs und die diesen Ablauf beeinflussenden Aktionen der teilnehmenden Dialogpartner beschrieben. Sowohl Wahrnehmung als auch Einleitung dieser Aktionen bilden demnach elementare Fähigkeiten, um innerhalb eines Gesprächs mehrere gemeinsame Themen und einen kohärenten Dialogablauf aufbauen zu können. Die automatische Identifikation dieser Aktionen und die Integration der daraus resultierenden Informationen in das konversationale Verhalten des Zielsystems bildet somit einen grundlegenden Aspekt bei der Konzeption künstlichen Themenbewusstseins für den Agenten Max.

Ein weiterer Schwerpunkt dieses Kapitel lag in der Darstellung des mentalen Lexikons, unserem internen Wörterbuch, welches unser gesamtes Wissen über Konzepte und ihre Zusammenhänge beinhaltet. Es bildet somit die Grundlage für die Fähigkeit, die in einem Gespräch angesprochenen Konzepte zu einem Thema zusammenzufassen. Menschliches Themenbewusstsein in laufenden Gesprächen kann daher nur unter kontinuierlichem Zugriff auf dieses Lexikon gelingen. Diese Erkenntnis liefert ein weiteres Kriterium für die Konzeption künstlichen Themenbewusstseins, deren vollständige Beschreibung in Kapitel 4 erfolgt.

3 Grundlagen künstlichen Themenbewusstseins

Obwohl die Erkennung und der Umgang mit Dialogthemen für uns Menschen eine Selbstverständlichkeit darstellt, stellt uns die Automatisierung dieser intuitiv ablaufenden Prozesse vor eine große Herausforderung. Gegeben sei folgender Dialog:

A: *„Ich gehe morgen zum Spiel, Bayern gegen Schalke. Magst du mitkommen?“*

B: *„Nein, tut mir leid. Ich mag keinen Fußball.“*

An sich ist es sehr einfach zu beurteilen, worüber sich die beiden Sprecher unterhalten: das Thema ist Sport. Aufgrund dieser Information ist es Menschen möglich, den mehrdeutigen Begriff „Spiel“ richtig aufzulösen und kohärente Sequenzen an Dialogbeiträgen zu erzeugen und auszutauschen. Wie kann dies automatisch, zum Beispiel bei einem künstlichen Dialogpartner, gelingen? Für eine Emulation menschlichen Themenbewusstseins in natürlichen Interaktionen mit künstlichen Systemen müssen zwei Hauptaufgaben automatisiert werden: Erstens, die Erkennung von thematischen Aktionen wie Themeneinführung, Themenwechsel, etc. in natürlichsprachlichen Dialogen und zweitens, die adäquate Integration dieses Themenwissens in die entsprechende Systemarchitektur. Während Letzteres stark von der Architektur des jeweiligen Dialogsystems abhängt, lässt sich die automatische Identifikation von Gesprächsthemen systemunabhängig realisieren. Je nach Anwendungsgebiet fallen Anforderungen und Umsetzung der hierfür umgesetzten Implementierung unterschiedlich aus, Grundvoraussetzung ist jedoch immer der Zugriff auf konzeptuelles Wissen, das heißt auf Wissen über Konzepte und ihre Zusammenhänge. Die Erfüllung dieser Voraussetzung bildet eine weitere, im Rahmen dieser Arbeit zu bewältigende Aufgabe.

Ziel dieses Kapitels ist es, elementare Methoden und unterschiedliche Ansätze für die für eine Implementierung künstlichen Themenbewusstseins erforderlichen Aufgaben aufzuzeigen sowie einen Überblick über den aktuellen Stand der Forschung zu ver-

mitteln. Demnach gliedert sich dieses Kapitel wie folgt: Abschnitt 3.1 beschäftigt sich mit der für eine automatisierte Themenerkennung erforderlichen maschinenlesbaren Darstellung konzeptuellen Wissens. Dabei wird insbesondere die Online-Enzyklopädie Wikipedia als adäquate Informationsquelle vorgestellt, welche das Wissen über einzelne Konzepte in Form von zahlreichen, natürlichsprachlichen Textbeschreibungen zur Verfügung stellt. Da jedoch die Inhalte maschinell nur schwer zugänglich sind und die Suche nach bestimmten Informationen erschweren, wird nach einer Lösung für einen schnellen Zugriff auf die in den Texten enthaltenen Informationen gesucht. Die Einführung der dazu erforderlichen texttechnologischen Methoden schließt diesen Abschnitt ab.

Anschließend, in Abschnitt 3.2, werden Grundlagen automatischer Themenerkennung vermittelt. Zu diesem Zweck werden zunächst die für die vorliegende Arbeit relevanten Techniken der textbasierten Informationsbeschaffung im Sinne des Information Retrieval dargestellt. Angewendet auf Wikipedia ermöglichen diese die Verwendung kollaborativ erstellter Konzeptbeschreibungen und -strukturen als maschinenlesbares Lexikon. Schließlich liefert Abschnitt 3.3 einen Überblick über den aktuellen Stand der Forschung hinsichtlich automatisierten Themenbewusstseins in der Mensch-Maschine-Interaktion. In diesem Zusammenhang werden bestehende natürlichsprachliche Schnittstellen und die jeweiligen Realisierungen künstlichen Themenbewusstseins beschrieben.

3.1 Konzeptwissen für Maschinen

Die für eine thematische Gesprächsführung grundlegende Fähigkeit, sprachliche Äußerungen zu verstehen, setzt die Kenntnis vieler Fakten und Zusammenhänge voraus, die der Mensch im Laufe seines Lebens als sogenanntes *Weltwissen* erlangt (Gängler & Wachsmuth, 1993). Folglich ist es notwendig, auch Dialogsystemen einen Zugriff auf ebensolches Wissen zu gewährleisten, um kohärente Dialoge zwischen Menschen und Maschinen zu ermöglichen. Dieser Abschnitt beschäftigt sich mit der maschinenlesbaren Darstellung konzeptuellen Wissens und der Herausforderung, dieses Wissen effizient zugänglich zu machen.

3.1.1 Wissensrepräsentation

McCarthy (1959) argumentierte als Erster, dass Maschinen Wissen benötigen, genau wie Menschen ihren Verstand, um intelligentes Verhalten zu simulieren (Ponzetto, 2010). Folglich stellt in vielen Bereichen der Informatik die Repräsentation und der Austausch von Wissen eine wesentliche Aufgabe dar. Ähnlich wie Menschen auf ihr internes Wissen zurückgreifen und Lehrbücher, Regelwerke und Lexika verwenden können, benötigen Maschinen eine Repräsentation der zugrunde liegenden Begriffe und Zusammenhänge (Hesse, 2002). Da die Inhalte menschlicher mentaler Lexika implizit und nur begrenzt formalisierbar und konkret darstellbar sind, bedienen sich Softwareentwickler zu diesem Zweck der sogenannten *Wissensrepräsentation*, welche das benötigte Wissen in einer Form darstellt und strukturiert, die sowohl den Bedürfnissen der Menschen als auch den Anforderungen der technologischen Umsetzbarkeit genügt (Kienreich & Strohmaier, 2006). Das Gebiet der Wissensrepräsentation umfasst dabei zwei Aufgaben: Die Modellierung der benötigten Wissensbereiche und die Entwicklung von Repräsentationsformalisten, die dieses Wissen maschinenlesbar darstellen (Meyer-Fujara, Puppe & Wachsmuth, 1995). Es existieren verschiedene Formalismen, die sich vor allem in ihrer Mächtigkeit, Fakten in deutlicher, präziser und natürlicher Art darzustellen und neue Fakten auf Grundlage des bereits repräsentierten Wissens ableiten zu können, unterscheiden.

Repräsentationsformalisten

Im Folgenden werden unterschiedliche Repräsentationsformalisten vorgestellt. Ihnen allen ist gemein, dass sie auf einer *Graph-basierten* Darstellung des enthaltenen Wissens aufbauen. Ein **Graph** besteht gemeinhin aus einer Menge an Objekten, *Knoten* genannt, und einer Menge an Verbindungen in Form von *Kanten*, die die Objekte miteinander in Beziehung setzen und so einen geometrischen Raum aufspannen (Widdows, 2004). Um ein gemeinsames Verständnis Graph-basierter Wissensrepräsentation zu schaffen, bedienen wir uns der Darstellung Ponzettos (2010):

Sei $V = I \cup C$ eine Menge an Konzepten, bestehend aus einer Menge an Individuen I (z. B. Angela Merkel) und einer Menge an Klassen C , die Individuen mit gleichen Eigenschaften zusammenfassen (z. B. Bundeskanzler von Deutschland). Ferner sei R eine Menge an semantischen Relationen zwischen den Individuen und Klassen, die in V enthalten sind (z.B. *instance-of*, *isa*, *part-of*, usw.), die möglicherweise eine leere (un-

spezifizierte) semantische Relation r_{null} enthält. Die innerhalb dieser Arbeit betrachteten Wissensressourcen können als beschriftete, gerichtete Graphen $G = \langle V_G, E_G \rangle$ betrachtet werden, wobei $V_G \subseteq V$ die Menge der Knoten und $E_G \subseteq V_G \times R \times V_G$ die Menge der Kanten in G ist. Da die Kanten innerhalb eines gerichteten Graphen eine Richtung besitzen, sind die Elemente in E_G geordnete Tripel und für jeden Knoten kann zwischen eingehende und abgehende Kanten unterschieden werden. Die Menge an Knoten eines Graphen, die keine eingehenden Kanten besitzen, werden *Wurzelknoten* genannt. Wenn G gegeben ist, werden die Wurzeln definiert als $V_0 = \{v_0 | v_0 \in V_G \wedge \nexists v, r : \langle v, r, v_0 \rangle \in E_G\}$. Zwei Kanten sind benachbart, wenn sie einen gemeinsamen Knoten haben. Eine Sequenz benachbarter Kanten in einem Graphen wird als *Pfad* bezeichnet, wenn sich alle besuchten Knoten voneinander unterscheiden. Dabei definiert sich die Länge eines Pfades durch die Anzahl seiner Kanten. Unterscheiden sich die benachbarten Knoten nicht alle voneinander, so beschreibt die Sequenz einen *Zyklus*. Gerichtete Graphen, die Zyklen enthalten, werden als *zyklische Graphen* definiert, andernfalls wird der Graph als *azyklisch* bezeichnet.

Kontrolliertes Vokabular Ein kontrolliertes Vokabular enthält eine Liste an Begriffen, die eindeutig und möglichst redundanzfrei definiert sind (National Information Standards Organization, 2005). Das heißt, kontrollierte Vokabulare beinhalten keine lexikalischen Mehrdeutigkeiten und vermeiden das Auftreten von Synonymen, indem jedem Konzept genau eine, oder zumindest eine bevorzugte, Bezeichnung zugewiesen wird (Kienreich & Strohmaier, 2006). Beispiele für kontrollierte Vokabulare sind Indizes, Glossare und Schlagwortregister. Kienreich und Strohmaier weisen auf die Unzulänglichkeit dieses Formalismus' für komplexe Aufgabenstellungen der Wissensmodellierung hin, da keine Strukturen abgebildet werden können und Kontextabhängigkeiten von Bezeichnungen nicht reflektiert werden.

Taxonomie Die in einer Taxonomie enthaltenen Konzepte werden in einer meist hierarchisch organisierten Hyperonymie-Struktur klassifiziert, das heißt in Klassen gleicher Eigenschaften geordnet. Dies geschieht anhand von *Subsumtionsrelationen*, die *isa-Beziehungen* (deutsch: *ist ein(e)*-Beziehungen) zwischen zwei Konzepten spezifizieren. Ponzetto (2010) zufolge besteht eine Subsumtionsrelation zwischen zwei Konzepten c_1 und c_2 genau dann, wenn alle Elemente von c_1 auch Elemente von c_2 sind. Dargestellt wird diese Relation durch eine gerichtete Kante von c_1 zu c_2 . Auf diese Weise

lässt sich zum Beispiel *Hund isa Tier* ausdrücken. Ebenso kann eine Taxonomie *Instanzierungsrelationen* enthalten, die sogenannte *instance-of*-Beziehungen (deutsch: *ist Instanz von*-Beziehungen) als Spezialfall der Subsumption definieren, wenn c_1 auf ein Individuum verweist. Die Spezifikation Angela Merkels als deutsche Bundeskanzlerin lässt sich beispielsweise durch *Angela Merkel instance-of Bundeskanzler von Deutschland* darstellen. Ein Beispiel für eine Taxonomie, welche eine zentrale Rolle innerhalb dieser Arbeit einnimmt, ist das Kategoriensystem von Wikipedia (s. Kapitel 3.1.2).

Semantisches Netz Ein semantisches Netz erweitert die Mächtigkeit einer Taxonomie, indem es die Spezifikation weiterer qualifizierter Relationen zwischen den Konzepten, wie die von Synonymrelationen oder kausalen Zusammenhängen, erlaubt. Damit bildet es die Grundlage für viele moderne Anwendungen wie des *Natural Language Processing* oder des *Information Retrieval* (s. Abschnitt 3.2.1). Eine in diesem Formalismus repräsentierte Wissensansammlung ist *WordNet*, welche im nächsten Abschnitt näher vorgestellt wird.

Ontologie Im Wesentlichen auf semantischen Netzen beruhend, diese jedoch in einigen Punkten erweiternd, ist eine Ontologie eine formale Spezifikation einer Konzeptualisierung zum Aufbau eines gemeinsamen Verständnisses einer bestimmten Domäne. In der Praxis bestehen Ontologien aus Instanzen, Konzepten, Relationen und einer Menge an *Axiomen* (Aussagen), die auf diesen Relationen basieren und die Ableitung neuer Informationen ermöglichen (Ponzetto, 2010). Eine im Anschluss an diesem Abschnitt näher betrachtete Ontologie ist die Wissensdatenbank *Cyc*.

Repräsentation menschlicher Wissensstrukturen

Die Übertragung menschlicher Wissensstrukturen und -inhalte in ein maschinenlesbares Format kann *automatisch* oder *manuell* erfolgen. Für den manuellen Schritt werden menschliche Experten beauftragt, ihr Wissen bezüglich der gewünschten Domäne in Form von Konzepten und semantischen Relationen zu definieren. Abbildung 3.1 zeigt ein Beispiel für ein semantisches Netz, welches von Biologen zur Darstellung von Genen und ihren Eigenschaften im Rahmen der sogenannten *Gene Ontology*-Initiative¹ erstellt wurde. Zwei bekannte Projekte dieser Vorgehensweise sind *WordNet* und *Cyc*.

¹<http://www.geneontology.org/>

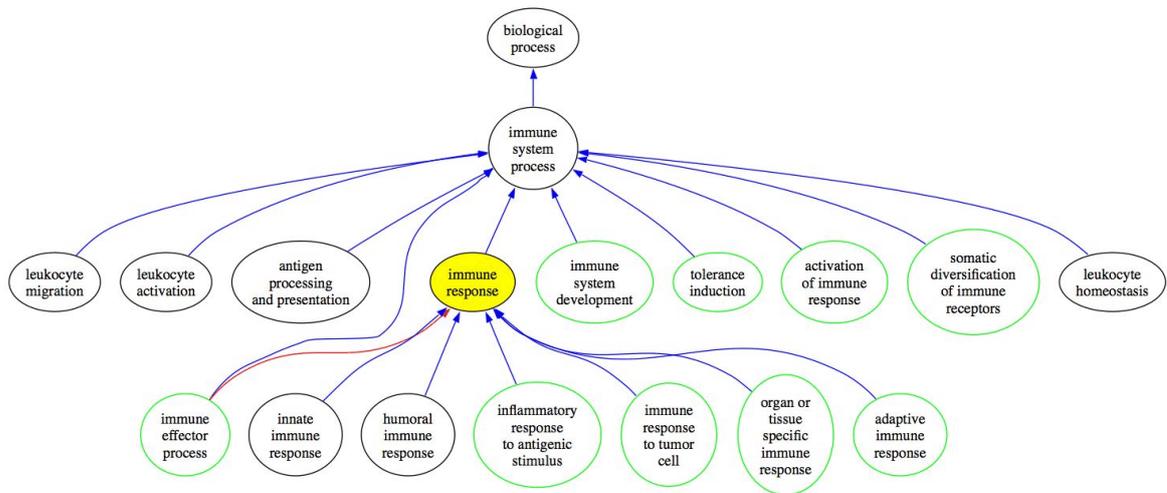


Abbildung 3.1: Ein im Rahmen der *Gene Ontology*-Initiative manuell erstelltes semantisches Netz zur Repräsentation von Prozessen des menschlichen Immunsystems. Quelle: <http://www.geneontology.org/GO.immunology.shtml>

Hervorgegangen aus der Idee des Psycholinguisten Georg A. Miller (1985), die (Un-)Menge lexikalischer Konzepte einer Sprache in einer hierarchischen, Netzwerk-artigen Struktur zu repräsentieren, stellt **WordNet** ein großes, manuell erstelltes semantisches Netz bereit, welches Wörter der englischen Sprache mit ähnlicher Bedeutung miteinander verknüpft (Fellbaum, 2010). So ist WordNet eine elektronische, lexikalische Datenbank, dessen Design von psycholinguistischen Theorien und den Strukturen menschlicher mentaler Lexika inspiriert wurde und welche lexikalische Informationen hinsichtlich ihrer Bedeutungen anstatt ihrer Bezeichnungen organisiert. Insofern gleicht WordNet eher einem kontrollierten Vokabular als einem Wörterbuch. Seit 1997 wird an der Universität Tübingen ein lexikalisch-semantisches Netz für die deutsche Sprache entwickelt, *GermaNet* genannt. Die Grundstruktur ist der von WordNet sehr ähnlich, wurde jedoch leicht modifiziert (zum Beispiel durch eine Kennzeichnung künstlich angelegter Konzepte, um diese von den weiteren Konzepten unterscheiden zu können) (Hamp & Feldweg, 1997). Für eine Nutzung von GermaNet stehen unterschiedliche Lizenzen zur Verfügung, unter anderem eine kostenlose für wissenschaftliche Zwecke.

Das von Douglas B. Lenat 1984 gegründete **Cyc**-Projekt² befasst sich mit der formalisierten Repräsentation menschlichen Alltagswissens in Form von Begriffen, Regeln und Zusammenhängen. Alle Inhalte werden als logische Aussagen in der auf der Prä-

²<http://cyc.com/cyc/technology/whaticyc>

dikatenlogik aufbauenden Ontologiesprache *CycL* definiert. Die Wissensbasis umfasst manuell eingepflegte Daten in Form von nahezu 500.000 Begriffen, 25.000 Relationsarten und mehr als fünf Millionen Fakten, die Aspekte wie „wenn ein Glas umkippt, fließt der Inhalt hinaus“ oder „wenn ein Objekt fliegt, berührt es nicht den Boden“ ausdrücken. Eine zusätzliche Inferenzmaschine erlaubt zudem die Ableitung weiteren Wissens, sodass die Wissensbasis von Cyc ein sehr hohes Maß menschlichen Wissens bereitstellt. Geht man allerdings von Schuberts (2006) Auffassung aus, dass das absolute Minimum für ein domänenunabhängiges, natürliches Gespräch zwischen Mensch und Maschine das Hundert- oder sogar Tausendfache an repräsentiertem Wissen erfordert, liegen noch einige Jahrzehnte Arbeit vor den Mitarbeitern des Projekts, um diesen Wissensstand zu erreichen.

Obschon WordNet, GermaNet und Cyc das jeweils enthaltene Wissen in sowohl wohlstrukturierter als auch -organisierter Form repräsentieren, veranschaulichen diese Projekte die Schwierigkeiten der manuell erstellten Wissensbasen hinsichtlich ihrer Verwendung in realen, domänenunabhängigen Künstliche Intelligenz- und Sprachverarbeitungsanwendungen. Nach Ponzetto (2010) liegen die Probleme hauptsächlich in der Abdeckung aller Wissensbereiche, in der Erhaltung der Objektivität bei der Auswahl und Darstellung von Informationen und in der Gewährleistung der Aktualität der präsentierten Daten. Der Aufbau und die regelmäßige Wartung von manuell erstellten Wissensrepräsentationen ist daher sehr aufwendig, sodass in den letzten Jahrzehnten ein neuer Forschungsbereich entstanden ist, der den Fokus auf automatische Wissensakquisition legt: der Bereich der **Informationsextrahierung** (engl. *information extraction*).

Die Informationsextrahierung (IE) beschäftigt sich mit maschinellen Verfahren zur Interpretation und Extrahierung strukturierter Informationen aus unstrukturierten, in natürlicher Sprache verfassten Texten. Techniken automatischer Sprachverarbeitung erlauben es, diese Texte vorverarbeiten zu können, sie mit sprachlichem Wissen anzureichern, ihnen die gewünschten Informationen zu entnehmen und letztere in strukturierter Form zurück zu liefern. Die dafür erforderlichen Schritte lassen sich auf einen Prozess abbilden, der zunächst die syntaktische Struktur des Textes erkennt, darauf aufbauend seinen Inhalt analysiert und letztendlich eine gegebene Struktur mit Informationen aus dem Text füllt (Dengel, 2012). Hierfür werden Sprachverarbeitungstechniken wie die Zuordnung von Textbestandteilen zu Wortarten (s. *Part-of-Speech*

Tagging, Abschnitt 3.1.4) oder die Erkennung und Klassifikation von Eigennamen und Objekten (s. *Named Entity Recognition*) eingesetzt.

Einsatz fanden diese Verfahren schon früh, beispielsweise im LILOG Projekt (Herzog & Rollinger, 1991) zur automatischen Erstellung touristischer Auskunftssysteme. Aktuell werden sie insbesondere in Suchmaschinen wie *Google* verwendet, aber beispielsweise auch in digitalen Bibliotheken, um die Suche nach der passenden Literatur zu unterstützen. Darüber hinaus können sie zum automatischen Aufbau *domänenunabhängiger* Wissensbasen genutzt werden. Voraussetzung dafür ist ein maschineller Zugriff auf umfangreiche Ansammlungen schriftlicher Texte. Das World Wide Web stellt einen enormen Bestand kostenloser Textdaten unterschiedlichster Themenbereiche bereit und hat sich gerade in den letzten Jahren als bedeutende Informationsquelle zur automatischen Akquisition von Wissen für die Entwicklung intelligenter Systeme etabliert (Ponzetto, 2010). Eine vollautomatisierte Akquisition taxonomischen Wissens auf Basis von online verfügbaren Daten ist jedoch nicht ohne Weiteres realisierbar, da die dezentralisierte, unkontrollierte Veröffentlichung von Web-Inhalten gewaltige Heterogenität der Webseiten und der enthaltenen Informationen zur Folge hat (Manning, Raghavan & Schütze, 2008). Demgegenüber stellt die Online-Enzyklopädie **Wikipedia** einen riesigen Speicher teilstrukturierter, lexikalischen Wissens in Form von zahlreichen Artikeln bereit. Die zusätzliche Klassifikation dieser Artikel in Ober- und Unterkategorien erlaubt die Extrahierung taxonomischer Wissensstrukturen, die für den Aufbau eines maschinenlesbaren mentalen Lexikons entscheidend sind. Nachfolgend wird eine ausführliche Beschreibung der Wikipedia und ihrer Spezifikationen gegeben.

3.1.2 Wikipedia

Als Resultat der Zusammenarbeit zahlreicher freiwilliger Autoren stellt Wikipedia eine Sammlung enzyklopädischen Wissens in einer Größenordnung bereit, wie es sie noch nie zuvor gegeben hat. Dabei bildet die im März 2001 gegründete deutsche Wikipedia mit aktuell über 1,4 Millionen Artikeln³ die zweitgrößte Version der in rund 285 Sprachen⁴ erhältlichen Enzyklopädie und verzeichnet bis heute ein konstantes Wachstum von durchschnittlich etwa 450 neuen Artikeln pro Tag⁵. Jeder dieser Artikel befasst

³http://de.wikipedia.org/wiki/Deutschsprachige_Wikipedia, Stand: 19. Oktober 2012

⁴<http://de.wikipedia.org/wiki/Wikipedia>, Stand: 19. Oktober 2012

⁵<http://de.wikipedia.org/wiki/Wikipedia:Wikipedistik/Wachstumsprognose>,
Stand: 19. Oktober 2012

sich mit der Beschreibung eines einzelnen Konzepts (s. Artikeldarstellung in Abbildung 3.2) und kann von jedem, ob anonym oder über ein Benutzerkonto, erstellt, korrigiert und erweitert werden. Darüber hinaus haben die ehrenamtlichen Autoren die Möglichkeit, den einzelnen Artikeln Kategorien zuzuordnen, die ebenfalls von den freiwilligen Autoren definiert und angelegt werden. Querverweise via *Hyperlinks* erlauben zudem die direkte Verknüpfung von Artikel- und Kategorienseiten und erleichtern Benutzern die Navigation. So bestehen die Artikel nicht nur aus natürlichsprachlichen, in Fließtexten verfassten Konzeptbeschreibungen, sondern enthalten darüber hinaus strukturierte Informationen. Die Inhalte der Wikipedia sind somit *teilstrukturiert*.

The image shows a screenshot of a Wikipedia article titled "Avatar Max". Several elements are highlighted with red boxes and labeled with red text and arrows:

- Artikelname:** A red box around the title "Avatar Max" is connected by a red line to the label "Artikelname".
- Hyperlinks zu anderen Artikeln:** A red box around the text "Stefan Kopp" and "Ulrike Wachsmuth" in the main text is connected by a red line to the label "Hyperlinks zu anderen Artikeln".
- Kategorien:** A red box around the category list "Kultur (Paderborn) Künstliche Intelligenz" at the bottom is connected by a red line to the label "Kategorien".

Other visible elements include the Wikipedia logo, navigation tabs (Artikel, Diskussion, Lesen, Bearbeiten, Versionsgeschichte), a search bar, a sidebar with navigation options, and a 3D avatar of a man in a purple shirt.

Abbildung 3.2: Wikipedia-Artikel über den virtuellen Agenten Max.

Das von Wikipedia umgesetzte kollaborative Verfahren der Lexikonerstellung erzielt zum einen eine Art *Objektivität* der enthaltenen Informationen und zum anderen eine beachtliche Qualität, die der der *Encyclopedia Britannica*, einer der ältesten Enzyklopädien der englischen Sprache, hinsichtlich der Korrektheit und der Abdeckung gleich-

kommt (Giles, 2005). Entsprechend groß ist der Erfolg des Web-basierten Lexikons. So steht Wikipedia laut Internetdienst *Alexa* auf Platz 6 der am häufigsten aufgerufenen Websites⁶, und die Suchmaschine *Google* liefert für eine Vielzahl an Suchbegriffen als passendsten Treffer den entsprechenden Wikipedia-Artikel zurück (Kohlenberg, 2006). Allerdings ist Wikipedia, wie jede andere Enzyklopädie auch, ungenau, nicht fehlerfrei und zudem häufigem Vandalismus ausgesetzt. Viégas, Wattenberg und Dave (2004) stellten jedoch anhand einer Visualisierung der Evolution von Artikeln (der sogenannten *history flow*-Visualisierung) fest, dass die aktive Wikipedia-Gemeinde die meisten dieser Defekte schnell und effektiv repariert. Dennoch führte die Administration der deutschen Wikipedia, entgegen dem allgemeinen Grundsatz Wikipedias, jede mögliche Aufbereitung direkt für alle sichtbar zu machen, ein System zur Wahrung der Qualität und der Vermeidung von Vandalismus ein, bei dem Änderungen der Artikel zuerst durch regelmäßige Autoren gesichtet werden müssen, bevor sie anschließend veröffentlicht werden.

3.1.3 Diskussion

Wie zu Beginn des Kapitels erwähnt, ist für die Realisierung einer automatischen Themenerkennung der Aufbau eines maschinenlesbaren, mentalen Lexikons erforderlich, der einen maschinellen Zugriff auf Wissen über Konzepte und ihre Zusammenhänge gewährleistet. Da die manuelle Erstellung eines solchen Konzeptwissens sehr aufwendig ist, wird nach einer geeigneten Informationsquelle gesucht, die dieses Wissen bereitstellt. Die vorgestellten Informationsquellen besitzen unterschiedliche Vor- und Nachteile, die in Tabelle 3.1 anschaulich zusammengefasst werden. Die Anordnung der einzelnen Kriterien spiegelt dabei eine Rangfolge wider, sodass das oberste Kriterium, die Forderung nach in deutscher Sprache verfassten Informationen, von äußerster Bedeutung für die Wahl der jeweiligen Informationsquelle ist, während das unterste Kriterium weniger ausschlaggebend für die Wahl ist.

Zusammengefasst bietet Wikipedia nicht nur einen Zugriff auf enzyklopädisches Wissen, dessen Erstellung und Pflege in den Händen freiwilliger Autoren liegen, sondern auch einen Zugang zu Beschreibungen und Kategorisierungen von Konzepten, welche von zahlreichen Mitwirkenden angelegt und editiert wurden und somit Einsichten in die allgemeine menschliche Auffassung konzeptueller Strukturen liefern. Diese im Kon-

⁶<http://www.alexa.com/topsites>, Stand: 19. Oktober 2012

Tabelle 3.1: Verschiedene Quellen konzeptuellen Wissens im Vergleich. Die einzelnen Kriterien sind hinsichtlich ihres Stellenwerts sortiert, sodass die oberen Kriterien ausschlaggebender sind als die unteren. Die mit * markierten Häkchen des Verfügbarkeitskriteriums weisen auf eine frei verfügbare, zum Teil reduzierte Forschungslizenz hin.

	WordNet	GermaNet	CYC	Wikipedia	Internet
deutsch	-	✓	✓	✓	✓
domänenunabhängig	✓	✓	✓	✓	✓
aktuell	-	-	-	✓	✓
objektiv	-	-	-	✓	-
strukturiert	✓	✓	✓	halb	-
frei verfügbar	✓*	✓*	✓*	✓	✓

text der *Web 2.0*-Ära bezeichnete *kollektive Intelligenz* (engl. *collective intelligence*, s. (Waltinger & Breuing, 2012)) bietet gegenüber den anderen vorgestellten Quellen repräsentierten Konzeptwissens die entscheidenden Vorteile der Aktualität und Objektivität der Informationen. Diese, sowie die weiteren, in Tabelle 3.1 dargestellten Vorteile liefern die Gründe für die Wahl Wikipedias als Grundlage für den Aufbau eines maschinenlesbaren, mentalen Lexikons, welches für die Realisierung des im Rahmen dieser Arbeit konzipierten Themenmodells benötigt wird. Um jedoch große Teile der Wikipedia überhaupt als Ressource für maschinelle Anwendungen nutzen zu können, wird eine Darstellung der Wikipedia benötigt, die einen effizienten Zugriff auf die teilstrukturierten Inhalte ermöglicht, da der Online-Zugriff für zeitkritische Anwendungen ungeeignet ist. Zu diesem Zweck bietet die *Wikimedia Foundation*, welche die Wikipedia betreibt, frei verfügbare Downloads aller Wikipedia-Inhalte unterschiedlichster Sprachen in XML-ähnlicher Notation an. Eine wie im nachfolgenden Abschnitt 3.1.4 dargestellte Überführung dieser Daten in ein geeignetes Format unterstützt und beschleunigt die Suche nach bestimmten Informationen.

Eine weitere Möglichkeit, maschinell auf die Inhalte der Wikipedia zuzugreifen, bietet **DBpedia** (Bizer et al., 2009), welche den strukturierten Teil der Informationen extrahiert und in einer Ontologie (s. Abschnitt 3.1.1) frei zur Verfügung stellt. Die extrahierten Daten sind in *RDF*, dem *Resource Description Framework* (Klyne & Carroll, 2004) repräsentiert, einer Spezifikation des W3C zur Beschreibung von Ressourcen und ihren Eigenschaften. Die deutsche Version der DBpedia wird für die im Rahmen des KnowCIT-Projekts entwickelte QA-Komponente verwendet. Aufgrund

Tabelle 3.2: Index mit drei verschiedenen Dokumenten und den Wörtern HUND, KATZE und MAUS. Ist ein Wort in einem der Dokumente enthalten, erhält der entsprechende Indexeintrag eine 1; wenn nicht, wird dies durch eine 0 verdeutlicht.

	HUND	KATZE	MAUS
Dokument 1	1	1	0
Dokument 2	0	1	1
Dokument 3	1	0	0

der irregulären Aktualisierung der Informationen (etwa ein mal pro Jahr) und der Tatsache, dass die semantische Darstellung der Daten keinen Mehrgewinn für die automatische Themenerkennung darstellt, wird die DBpedia im Folgenden nicht weiter betrachtet⁷.

3.1.4 Indexierung von Textinhalten

Die Indexierung von Texten zielt darauf ab, die jeweiligen Inhalte effizient und günstig erreichbar- und erschließbar zu machen. Hierzu repräsentiert ein **Index** die einzelnen Textdokumente durch deren enthaltenen Wörter (Krüger-Thielmann & Paijmans, 2004). Tabelle 3.2 zeigt einen Beispielindex für drei Dokumente und drei darin vorkommende Wörter.

Bei der Indexierung kommen in der Regel verschiedene Verfahren zum Einsatz, um eine Auswahl der zu indexierenden Wörter, der sogenannten *Indexterme*, zu gewährleisten, die das jeweilige Dokument optimal repräsentieren. In einem ersten Vorverarbeitungsschritt wird der jeweils betrachtete Text in bedeutsame Bestandteile zerlegt. Die *Textsegmentierung* (engl. *text segmentation*) kann dabei auf verschiedenen Ebenen stattfinden, sodass sich Texte in Kapitel, Abschnitte, Paragraphen, Sätze, Wörter und sogar in Silben und Phoneme gliedern lassen (Feldman & Sanger, 2007). Die Segmentierung auf Wort- und Satzebene wird als **Tokenisierung** (engl. *tokenization*) bezeichnet. Je nach Form der betrachteten Texte bietet sich eine Tokenisierung anhand unterschiedlicher Trennzeichen wie Satz- oder Leerzeichen an. Die konkrete Umsetzung im Rahmen der vorliegenden Arbeit wird in Kapitel 5 erläutert.

In natürlichsprachlichen Texten sind meistens Wörter enthalten, die als Indexterme ungeeignet sind, da sie sehr häufig verwendet werden und somit nicht von Relevanz

⁷Interessierte Leser seien auf <http://de.dbpedia.org> verwiesen.

für die einzelnen Dokumente sind. Dazu gehören unter anderem Funktionswörter wie Artikel oder Präpositionen. Anhand einer sprachspezifischen **Stoppwortliste** können diese hochfrequenten, unbedeutenden Wörter identifiziert und von der Indexierung ausgeschlossen werden.

Eine weitere Möglichkeit, weniger charakteristische Wörter nicht mit in dem Index aufzunehmen, ist die Vereinigung ähnlicher Indexterme. Eine Dokumentensammlung enthält oft mehrere Varianten, sogenannte *Flexionen*, eines einzelnen Terms, um dessen jeweilige grammatikalische Funktion innerhalb eines Satzes zum Ausdruck zu bringen. Die Wörter „definieren“, „definierte“ und „Definition“ lassen sich zum Beispiel auf denselben Wortstamm zurückführen, obwohl sie innerhalb eines Satzes ganz unterschiedliche Funktionen einnehmen. Um solch ähnliche Wörter in einem einzigen Indexterm zusammenzufassen, werden die verschiedenen morphologischen Formen eines Wortes auf einen gemeinsamen Stamm reduziert. Für diesen Prozess stehen zwei Verfahren zur Verfügung (Krüger-Thielmann & Pajjmans, 2004):

- **Stammformreduktion** Reduzierung des betrachteten Wortes zu seiner Stammform durch Entfernung von Flexionsendungen und Suffixen. Das Ergebnis muss kein in der jeweiligen Sprache existierendes Wort sein.
Bsp.: Freiheiten → frei
- **Lemmatisierung** Rückführung der flektierten Wortform auf dessen Grundform (*Lemma*) durch Entfernung von Flexionsendungen, Umlauten und graphemischen Veränderungen.
Bsp.: Freiheiten → Freiheit; Bsp.(graphemisch): Nougat → Nugat

Durch beide genannten Verfahren wird die Anzahl unterschiedlicher Wörter innerhalb eines Textkorpus reduziert und dadurch das relative Vorkommen einzelner Wörter erhöht (Manning & Schütze, 1999). Darüber hinaus reduziert sich der Speicherbedarf des resultierenden Indexes aufgrund weniger Redundanzen. Ein großer Nachteil dieser Verfahren ist, dass sie wertvolle Informationen entfernen, die bei der Reduzierung der Wörter auf ihre Stamm- oder Grundformen verloren gehen (Jurafsky & Martin, 2009).

In einigen Anwendungen ist es sinnvoll, nur bestimmte Wortkategorien bei der Indexierung zu beachten. Ein Beispiel hierfür sind Nomina, die als die bedeutungstragenden Elemente in Texten gelten (Baeza-Yates & Ribeiro-Neto, 1999). Um einzelnen

Komponenten (Wörtern) des Texts ihre grammatikalischen Kategorien (beispielsweise „Nomen“ oder „Verb“) zuzuordnen und auf diese Weise zu ihrer Disambiguierung (Auflösung möglicher Mehrdeutigkeiten) beizutragen, wird das sogenannte **Part-of-Speech (PoS) Tagging** durchgeführt. Dabei kommen maschinelle Verfahren zum Einsatz, die die einzelnen Wörter mit ihrer jeweiligen Wortklasse (engl. *part-of-speech*) auszeichnen. Dieser Prozess muss *kontextabhängig* durchgeführt werden, da verschiedene Wortklassen in gleichen Wortformen verwendet werden können. In den Sätzen „Er verteidigte die Burg.“ und „Die verteidigte Burg fiel.“ wird die Wortform „verteidigte“ einmal als Verb und einmal als Adjektiv verwendet. Dies lässt sich nicht von der Wortform an sich, sondern nur durch die zusätzliche Betrachtung des Kontexts, das heißt der jeweils angrenzenden Satzkomponenten, ableiten (Heyer, Quasthoff & Wittig, 2006).

Für das Tagging-Problem wurden viele unterschiedliche Ansätze entwickelt, die sich in zwei Klassen einteilen lassen: regelbasierte und stochastische Tagger (Jurafsky & Martin, 2009). Regelbasierte Tagger beinhalten eine große Datenbank manuell erstellter Disambiguierungsregeln, die zum Beispiel spezifizieren, dass ein bestimmtes mehrdeutiges Wort eher ein Adjektiv als ein Verb ist, wenn das nachfolgende Wort ein Nomen ist. Stochastische Verfahren hingegen setzen auf Wahrscheinlichkeiten für die einzelnen Wortklassen auf, die auf Grundlage von Trainingsdaten maschinell berechnet werden. Beispiele für beide Ansätze liefern Jurafsky und Martin (2009, Kap. 5).

Die unterschiedlichen Vorverarbeitungsschritte ergeben eine Liste von Dokumenten und ihren jeweiligen enthaltenen Schlüsselwörtern, die anschließend in einem Index abgelegt werden (ein Beispiel für die konkrete Umsetzung liefert Kapitel 5). Um einen noch effizienteren Zugriff auf die Indexterme zu erhalten, werden die ursprünglich nach den Dokumenten geordneten Listen in Listen für die Schlüsselwörter transformiert (Krüger-Thielmann & Pajmans, 2004). Innerhalb dieses *invertierten Index* (engl. *inverted index*) verweist demnach jeder Indexterm auf die Dokumente, in denen er enthalten ist. Die Indexstruktur bietet auf diese Weise die Grundlage für maximale Performanz bei minimaler Nutzung der Ressourcen (McCandless, Hatcher & Gospodnetić, 2010).

3.2 Automatische Themenerkennung in Texten

Mit der Indexierung von Wikipedia-Inhalten wurde die erste Grundlage für ein automatisiertes Themenbewusstsein künstlicher Gesprächspartner geschaffen: der maschinelle Zugriff auf konzeptuelles Wissen. Die nächste Herausforderung ist die Automatisierung der Erkennung angesprochener Themen und der den thematischen Ablauf des Dialogs beeinflussenden Aktionen auf Basis dieses repräsentierten Konzeptwissens. Zur Bewältigung dieser Aufgabe müssen zunächst die aufeinander folgenden Dialogbeiträge auf thematische Ähnlichkeit überprüft werden. Anschließend gilt es diese Blöcke thematisch zu klassifizieren und zu benennen. Dementsprechend stellt dieser Abschnitt zunächst die für die vorliegende Arbeit relevanten technischen Grundlagen für die Bestimmung der Ähnlichkeit von Texten vor. Im Anschluss wird der Prozess der *Textklassifikation* eingeführt und in diesem Zusammenhang Wikipedia als Quelle für vordefinierte Klassifikationskategorien präsentiert. Der Abschnitt endet mit der Vorstellung der im Rahmen des *TDT*-Projekts definierten Aufgaben, die es für eine automatische Extrahierung thematischer Strukturen aus Radionachrichten zu bewältigen gilt und dessen Einteilung in einzelne Komponenten als Grundlage für das im Rahmen der vorliegenden Arbeit konzipierten Themenmodells dient.

3.2.1 Textähnlichkeit

Das Fachgebiet der Informationserschließung, das sogenannte **Information Retrieval**, zielt auf die Entwicklung von Modellen und Algorithmen ab, die die Suche nach Informationen in großen Datenbeständen unterstützen (Manning & Schütze, 1999) (beispielsweise die Stichwortsuche in Suchmaschinen). Anfangs als sehr eingeschränktes Interessensgebiet für Bibliothekare und Informationsexperten angesehen, hat das Gebiet des IR insbesondere seit Einführung des World Wide Web einen enormen Aufschwung erlebt und neue Forschungsschwerpunkte und Technologien im Bereich der textbasierten Suche hervorgebracht (Baeza-Yates & Ribeiro-Neto, 1999). Das *Retrieval*, also das Auffinden des Dokuments mit der gewünschten Information, wird dabei zumeist durch einen Vergleich des Anfragetexts mit den zuvor indexierten Textdokumenten realisiert. Das Ergebnis ist eine bestenfalls geordnete Liste von Dokumenten, wobei die Reihenfolge der Dokumente die ermittelte Ähnlichkeit zwischen Anfrage und Dokument widerspiegelt (Krüger-Thielmann & Paijmans, 2004). Gelingen kann dies

mit Hilfe unterschiedlicher Modelle. Zu den klassischen Modellen zählen das *Boolesche Modell*, das *Probabilistische Modell* und das *Vektorraummodell*. Da das Boolesche Modell aufgrund binärer Bewertung keine Rangordnung der Ergebnisse unterstützt und das Probabilistische Modell auf Rückmeldungen zur Gewinnung von Relevanzwerten angewiesen ist, werden diese beiden Ansätze im Rahmen der vorliegenden Arbeit nicht weiter betrachtet⁸.

Das Vektorraummodell

Eines der verbreitetsten Modelle für Information Retrieval ist das **Vektorraummodell** (engl. *Vector Space Model*) (Salton, Wong & Yang, 1975). Es stellt die indexierten Dokumente als Elemente eines Vektorraums dar, dessen Raumdimensionen je einem in der betrachteten Dokumentensammlung enthaltenen Term entsprechen (Manning & Schütze, 1999). Da nicht alle Terme gleich bedeutend für das jeweilige Dokument sind, werden diesen numerische Werte, sogenannte *Gewichte* (engl. *weights*), zugeordnet, welche die Relevanz der einzelnen Indexterme für das betrachtete Dokument widerspiegeln (Baeza-Yates & Ribeiro-Neto, 1999). Sei $T = \{t_1, t_2, \dots, t_n\}$ die Menge aller Indexterme und d_j ein Dokument. Jedem Term $t_i \in T$ wird ein Gewicht $w_{i,j} > 0$ zugeordnet, das die Bedeutung von t_i für das Dokument d_j darstellt. Für einen Indexterm, welcher nicht in d_j enthalten ist, gilt $w_{i,j} = 0$. Auf diese Weise lässt sich d_j als n -dimensionaler Vektor $\vec{d}_j = (w_{1,j}, w_{2,j}, \dots, w_{i,j}, \dots, w_{n,j})$ repräsentieren, wobei n die Anzahl aller Indexterme in T ist. Bezogen auf die in Tabelle 3.2 (s. Abschnitt 3.1.4) beschriebenen Dokumente ergibt sich für das erste Dokument der Vektor $\vec{d}_1 = (1, 1, 0)$, für das zweite der Vektor $\vec{d}_2 = (0, 1, 1)$ und Dokument 3 wird durch den Vektor $\vec{d}_3 = (1, 0, 0)$ repräsentiert. Abbildung 3.3 veranschaulicht die Positionierung dieser Vektoren in einem 3-dimensionalen Raum.

Durch die zusätzliche Übertragung des Anfragetexts in diesen hochdimensionalen Raum, können die Dokumente, die die angefragte Information enthalten, nun anhand eines Vergleichs der Anfrage- und Dokumentenvektoren ermittelt werden. Dabei werden die Dokumente, deren Vektoren der der Anfrage am ähnlichsten sind und die somit die höchste Übereinstimmung bezüglich der beinhalteten Terme aufweisen, identifiziert und entsprechend ihrer Relevanz geordnet. Die Vorteile des Vektorraummodells liegen dabei in der Möglichkeit, unterschiedliche Merkmale für die Termgewichtung heran-

⁸Eine ausführlichere Darstellung liefern Baeza-Yates und Ribeiro-Neto (1999).

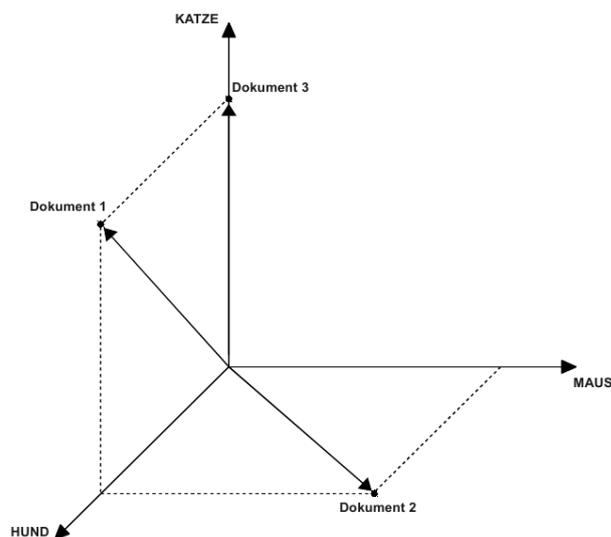


Abbildung 3.3: 3-dimensionaler Vektorraum für die Dokumente und Schlüsselwörter aus Tabelle 3.2 (in Anlehnung an (Krüger-Thielmann & Paijmans, 2004, S. 363)).

ziehen und auf eine Vielfalt an geeigneten Ähnlichkeitsmaßen zurückgreifen zu können (Krüger-Thielmann & Paijmans, 2004). Beide Aspekte werden im Folgenden näher erläutert.

Termgewichtung Die in Abbildung 3.3 dargestellte Gewichtung der Terme bringt lediglich zum Ausdruck, ob ein Term innerhalb eines Dokuments vorkommt oder nicht. Die Relevanz eines Dokuments hinsichtlich der gesuchten Information wird somit auf eine *binäre* Eigenschaft reduziert. Die Stärke des Vektorraummodells liegt jedoch in der Möglichkeit, reelle Zahlen für die Gewichtung der Terme zu verwenden und auf diese Weise eine Rangfolge der für eine bestimmte Suchanfrage relevanten Dokumente in der Ergebnisliste zu gewährleisten (Krüger-Thielmann & Paijmans, 2004). Unter der Annahme, dass Terme, die häufiger innerhalb eines Dokuments genannt werden, repräsentativer für genau dieses Dokument sind, kann eine differenziertere Gewichtung zum Beispiel durch die Berechnung des Termvorkommens gelingen. Die Anzahl der Vorkommnisse eines Terms t_i in einem Dokument d_j wird hierbei mit der sogenannten **Termfrequenz** $tf_{i,j}$ (engl. *term frequency*) erfasst, wobei der Wert in der Regel normiert wird, um die Länge des jeweiligen Dokuments zu berücksichtigen. Die

Tabelle 3.3: Bewährte Häufigkeiten für die Gewichtung von Termen in Dokumenten.

Häufigkeit	Symbol	Bedeutung
Termfrequenz	$tf_{i,j}$	Anzahl der Vorkommnisse von t_i in d_j
Dokumentfrequenz	df_i	Anzahl der Dokumente, in denen t_i vorkommt

normalisierte Termfrequenz ist somit definiert als

$$tf_{i,j} = \frac{freq_{i,j}}{\max_l freq_{l,j}} \quad (3.1)$$

wobei $freq_{i,j}$ die Anzahl von t_i in d_j und $\max_l freq_{l,j}$ die Anzahl des Terms mit dem höchsten Vorkommnis in d_j .

Bei der lediglich auf Termfrequenz basierenden Gewichtung wird Termen, die im Allgemeinen verbreiteter, dadurch aber weniger repräsentativ sind als andere, unter Umständen ein zu hohes Gewicht gegeben, da sie entsprechend oft in den Dokumenten vorkommen (Adams & Martel, 2010). Aus diesem Grund wird in den meisten IR-Systemen neben der Termfrequenz eine weitere Häufigkeit bei dem Gewichtungsverfahren beachtet: die **Dokumentfrequenz** df_i (engl. *document frequency*), die die Anzahl der Dokumente angibt, in denen der betrachtete Term t_i enthalten ist. Mit dem Ziel, Terme, die in wenigen Dokumenten vorkommen und deren Inhalte entsprechend besser charakterisieren, höher zu gewichten als solche, die in vielen Dokumenten enthalten sind, bedarf es einer logarithmischen Skalierung der Dokumentfrequenz. Die resultierende *inverse* Dokumentfrequenz

$$idf_i = \log \frac{N}{df_i} \quad (3.2)$$

überträgt Wörtern, die nur in einem einzigen Dokument vorkommen, das volle Gewicht, während Terme, die in allen Dokumenten der Sammlung vorkommen, ein Gewicht von 0 erhalten.

Das wohl bekannteste Gewichtsmaß im Bereich des Information Retrieval, das **tf-idf-Maß**, kombiniert die genannten, in Tabelle 3.3 noch einmal skizzierten Häufigkeiten in Form eines Produkts:

$$tfidf_{i,j} = tf_{i,j} \cdot idf_i = \frac{freq_{i,j}}{\max_l freq_{l,j}} \cdot \log \frac{N}{df_i} \quad (3.3)$$

wobei N die Gesamtzahl der Dokumente ist. Das Gewicht eines Terms für ein bestimmtes Dokument ist bei diesem Maß demnach besonders hoch, wenn dieser oft in dem Dokument selbst enthalten ist, über die gesamte Dokumentensammlung verteilt jedoch eine niedrige Frequenz aufweist.

Ähnlichkeitsmaße Im Zusammenhang mit dem Vektorraummodell wurden zahlreiche Funktionen zur Bestimmung der Ähnlichkeit zwischen der Suchanfrage und den einzelnen Dokumenten definiert. Zur Veranschaulichung des wohl bekanntesten Ähnlichkeitsmaßes, der **Kosinus-Ähnlichkeit** (engl. *cosine similarity*), dient folgendes Beispiel nach Manning und Schütze (1999): Gegeben sei ein wie in Abbildung 3.4 dargestellter zwei-dimensionaler Vektorraum, dessen Dimensionen den Indextermen „Auto“ und „Versicherung“ entsprechen. In diesem Raum sind drei Dokumente d_1 , d_2

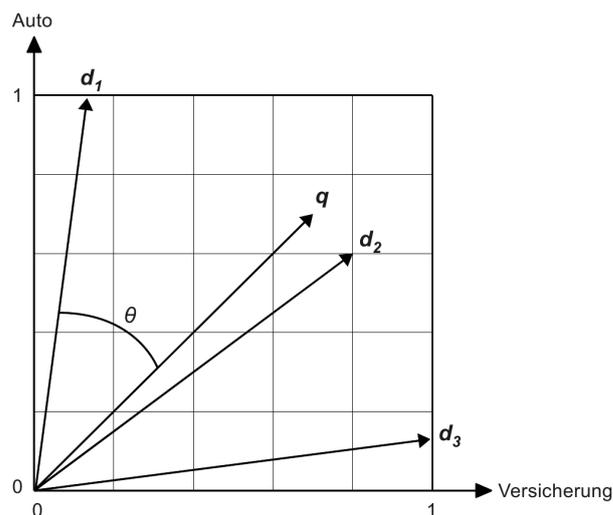


Abbildung 3.4: 2-dimensionaler Vektorraum mit Vektoren für drei Dokumente (d_1 , d_2 und d_3 und einer Informationsanfrage q (in Anlehnung an (Manning & Schütze, 1999, S. 540)).

und d_3 durch die Vektoren $\vec{d}_1 = (0,13 | 0,99)$, $\vec{d}_2 = (0,8 | 0,6)$ und $\vec{d}_3 = (0,99 | 0,13)$ sowie eine Suchanfrage q durch den Vektor $\vec{q} = (0,71 | 0,71)$ repräsentiert. Die Ähnlichkeit zwischen der Anfrage und einem der Dokumente lässt sich nun anhand des Kosinus des eingeschlossenen Winkels Θ zwischen den jeweiligen Vektoren bestimmen. Die Kosinus-Ähnlichkeit zwischen einem Dokument d_j und einer Informationsanfrage

q ist dabei definiert als

$$\text{sim}(d_j, q) = \frac{\vec{d}_j \bullet \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^n w_{i,j} \cdot w_{i,q}}{\sqrt{\sum_{i=1}^n (w_{i,j})^2 \cdot \sum_{i=1}^n (w_{i,q})^2}} \quad (3.4)$$

wobei $\vec{d}_j \bullet \vec{q}$ das *Skalarprodukt* und $|\vec{d}_j| \cdot |\vec{q}|$ die *euklidische Länge* der Vektoren $\vec{d}_j = (w_{1,j}, w_{2,j}, \dots, w_{n,j})$ und $\vec{q} = (w_{1,q}, w_{2,q}, \dots, w_{n,q})$ darstellen (Manning et al., 2008). Auf diese Weise wird die Termkorrelation zwischen den Vektoren berechnet, welche durch die euklidische Länge der beiden Vektoren dividiert wird. Letzteres ist erforderlich, um die Längen der einzelnen Vektoren, und damit die Länge der jeweiligen Dokumente, zu normieren (Manning & Schütze, 1999).

Die Herausforderung, das Dokument d_i zu finden, welches der Informationsanfrage q am meisten entspricht, reduziert sich somit auf die Ermittlung der höchsten Kosinus-Ähnlichkeit zwischen den entsprechenden Vektoren. Da die einzelnen Termgewichte nicht negativ sein können ($w_{i,j} \geq 0$), bewegt sich die Ähnlichkeit dabei zwischen $[0, 1]$. Die Dokumente werden somit gemäß ihres Ähnlichkeitsgrads zu der Anfrage klassifiziert. Bezogen auf das genannte Beispiel, lässt sich die größte Übereinstimmung zwischen q und d_2 berechnen, sodass die Wahrscheinlichkeit, die in q angeforderte Information in einem der Dokumente zu finden, für d_2 am größten ist.

Wikipedia-basierte Bestimmung der Textähnlichkeit

Mit Hilfe der Vektorraummodell-Repräsentation gelang es Gabrilovich und Markovitch (2007), durch einen Email-Service einer australischen Rundfunkgesellschaft verschickte Nachrichten auf Grundlage von Wikipedia miteinander zu vergleichen und die Ähnlichkeiten zu bestimmen. Die dafür entwickelte Methode, *Explicit Semantic Analysis* (ESA) genannt, basiert auf einem invertierten Index an Schlüsselwörtern und ihren Vorkommnissen in den Wikipedia-Artikeln. Auf diese Weise lässt sich ein Vektorraum aufspannen, dessen Dimensionen den einzelnen Wikipedia-Artikeln entsprechen und in dem sich alle in diesen Artikeln enthaltenen Schlüsselwörter als Vektoren in diesem Raum darstellen lassen. Um nun die Nachrichten miteinander vergleichen zu können, wird jedes Wort einer einzelnen Nachricht als Vektor repräsentiert, welcher ein Produkt aus der *tfidf*-Darstellung dieses Wortes und seinem entsprechenden Eintrag in

dem invertierten Wikipedia-Index darstellt. Anschließend werden alle Vektoren der Wörter einer Nachricht durch Addition in einem einzelnen Vektor zusammengefasst, sodass dieser schließlich die gesamte Nachricht repräsentiert. Der thematische Zusammenhang unterschiedlicher Nachrichten lässt sich nun anhand der Kosinus-Ähnlichkeit ihrer Vektoren bestimmen.

Ponzetto (Ponzetto, 2010) erweiterte diesen Ansatz im Rahmen seiner Dissertation um die bis dahin unberücksichtigte, durch das Kategoriensystem der Wikipedia bereitgestellte hierarchische Struktur. Die Ähnlichkeit zwischen zwei Wörtern w_1 und w_2 wird hier zum einen anhand von Termübereinstimmung und zum anderen auf Grundlage von Pfadlängen bestimmt. Hierzu werden die Wörter in einem ersten Schritt auf Wikipedia-Artikel abgebildet und anschließend die Kategorien der ermittelten Artikel extrahiert. Die semantische Ähnlichkeit ergibt sich dann aus der Ähnlichkeit der beiden Artikel (Termübereinstimmung) und der Ähnlichkeit der Kategorien innerhalb der Kategorientaxonomie (Pfadlänge). Von besonderer Bedeutung ist die in Ponzettos Ansatz durchgeführte *Konzeptualisierung* durch Abbildung der Wörter auf Wikipedia-Artikel, da diese den Ausgangspunkt für eine Ähnlichkeitsbestimmung zwischen einzelnen Wörtern ermöglicht. Die Idee der Konzeptualisierung wird in das im Rahmen der vorliegenden Arbeit konzipierte Themenmodell übernommen, um thematische Ähnlichkeiten zwischen den in den Dialogbeiträgen enthaltenen Termen zu berechnen (s. Kapitel 4).

3.2.2 Textkategorisierung

Unter *Klassifikation* oder *Kategorisierung* versteht man gemeinhin die Zuordnung von Objekten zu zwei oder mehr Klassen oder Kategorien. Dementsprechend bezeichnet die **Textkategorisierung** die Zuordnung neuer Textdokumente zu einer Menge an vordefinierten Themenkategorien (im Gegensatz zum *Textclustering*, bei dem die Kategorien aus der zugrunde liegenden Textstruktur selbst abgeleitet werden (Heyer et al., 2006)). Ein Anwendungsbeispiel sind Filterprozesse, die schon früh eine einfache, thematische Vorsortierung eingehender Emails ermöglichten, unter anderem auf Grundlage von vordefinierten regulären Ausdrücken (z. B. *Elm Filter* (Taylor, 1986), *Procmail* (Van den Berg, 1994) oder *Ishmail* (Helfman & Isbell, 1995)) oder durch das Erlernen der vom jeweiligen Benutzer durchgeführten Aktionen (z. B. *Re:Agent* (Boone, 1998)). Verfahren der Textklassifikation werden zudem in Suchmaschinen eingesetzt, um the-

matisch verwandte Webseiten in gebündelter Form anzuzeigen (G. Rehm, 2004; Chen & Dumais, 2000).

Nach Feldman und Sanger (2007) lässt sich die Textkategorisierung als Funktion $F : D \times C \rightarrow \{0, 1\}$ beschreiben, wobei D die Menge aller Dokumente und C die Menge aller Kategorien ist. Der Wert von $F(d, c)$ ist 1, wenn das Dokument d der Kategorie c angehört, ansonsten ist der Wert 0. Da sich die Kategorien überschneiden dürfen, können einzelne Dokumente mehreren unterschiedlichen Kategorien angehören. Das Standardvorgehen für eine automatische Umsetzung der Funktion ist die Anwendung überwachten Maschinellen Lernens. Das bedeutet, der resultierende *Klassifizierer* (engl. *classifier*) lernt die thematische Einordnung von Dokumenten anhand von Beispielen (zum Beispiel *Support Vector Machines (SVM)* oder *Bayes-Klassifikator*; eine Übersicht hierzu liefern Feldman und Sanger (2007, Kap. 4)). Die Kategorisierung von Texten kann auch manuell vorgenommen werden, was zwar zu sehr guten Ergebnissen führt, mit steigender Anzahl an Dokumenten jedoch mit höherem Aufwand verbunden ist. Ein Beispiel für die manuelle Einordnung von Textdokumenten in thematische Kategorien liefert das Kategoriensystem von Wikipedia.

Das Kategoriensystem von Wikipedia

Durch die bereits in Abschnitt 3.1.2 angesprochene Möglichkeit, den einzelnen Artikeln ein oder mehr Kategorien zuzuordnen zu können, unterstützt Wikipedia eine thematische Klassifikation der in den Artikeln beschriebenen Konzepte. Zu diesem Zweck stehen den Autoren der deutschen Wikipedia derzeit mehr als 55.000 Kategorien⁹ zur Verfügung, die assoziative Bezüge zwischen Artikeln und Kategorien, aber auch zwischen Kategorien untereinander herstellen. So gehört der Artikel über Künstliche Intelligenz beispielsweise der Kategorie „Kognitionswissenschaft“ an, die wiederum der Oberkategorie „Interdisziplinäre Wissenschaft“ zugeordnet ist. Auf diese Weise kommt eine taxonomische Kategorienstruktur zustande, die gemeinhin als *Kategorienbaum* bezeichnet wird (s. Abbildung 3.5). In der Praxis ist diese Struktur jedoch nicht strikt hierarchisch modelliert, sondern enthält wie in Abbildung 3.6 dargestellte Schleifen (Ponzetto, 2010) (auch wenn diese laut Kategorisierungsrichtlinien Wikipedias von den Autoren vermieden werden sollten) sowie von der Gesamtstruktur vollständig losgelöste Kategorien. Beide Phänomene treten jedoch äußerst selten auf, sodass die

⁹Stand: Januar 2012

Kategorienstruktur im Jahr 2007 zum Beispiel 99,8% aller Kategorienknoten umfasste und lediglich sieben Schleifen enthielt (Zesch & Gurevych, 2007). Nichtsdestotrotz handelt es sich bei dem vermeintlichen Baum eher um einen gerichteten azyklischen Graphen, weshalb er im Folgenden als **Kategoriengraph** bezeichnet wird. Die Knoten dieses Graphen, die keine ausgehenden Kanten besitzen, repräsentieren die Wikipedia-Artikel, alle weiteren Knoten stellen die Kategorien dar und die Kanten veranschaulichen die kategorischen Zuordnungen der Artikel und Unterkategorien.

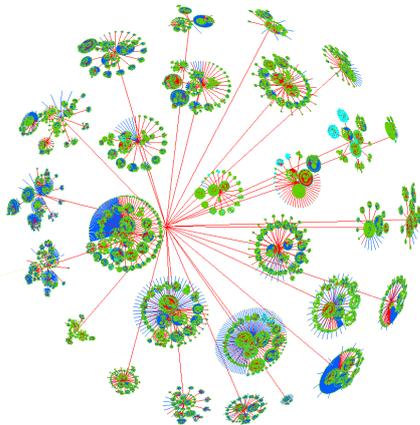


Abbildung 3.5: 3D-Darstellung des Kategoriensystems von Wikipedia, ausgerichtet auf die Hauptkategorie. Die roten Kanten stellen Verbindungen zu Kategorien (türkise Knoten) dar, die blauen veranschaulichen Hyperlinks zwischen Artikeln (grüne Knoten).

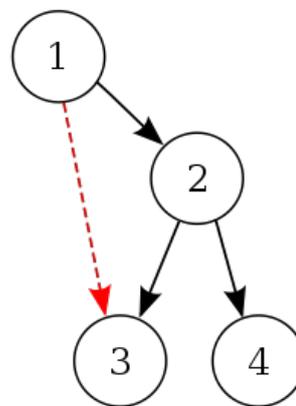


Abbildung 3.6: Eine Seite (Artikel oder Kategorie) sollte nicht gleichzeitig einer Kategorie und einer ihrer Unterkategorien zugeordnet werden, um Schleifen in dem Kategorienbaum zu vermeiden. Quelle: Wikipedia¹⁰.

Wie die Erstellung und Bearbeitung der Wikipedia-Artikel selbst, erfolgt auch die Einordnung der textbasierten Artikel in die Kategorien durch die freiwilligen Autoren. Der große Nachteil des enorm hohen Aufwands für eine manuelle Kategorisierung der über eine Million Artikel wird also auch hier durch die Kollaboration zahlreicher Freiwilliger ausgeglichen. Das Ergebnis ist eine thematisch organisierte Taxonomie, die aufgrund ihres Umfangs und ihrer Qualität die Grundlage vieler wissenschaftlicher Arbeiten und Ansätze bildet. So eignet sich das Kategoriensystem beispielsweise für die Verarbeitung von natürlicher Sprache (Zesch & Gurevych, 2007), zur Bestimmung semantischer Ähnlichkeiten (Chernov, Iofciu, Nejdil & Zhou, 2006) oder für die

¹⁰<http://de.wikipedia.org/wiki/Wikipedia:Kategorien>

Eingrenzung möglicher Antworten bei QA-Systemen (Vercoustre, Pehcevski & Thom, 2008; Waltinger, Breuing & Wachsmuth, 2012).

Textkategorisierung mit Hilfe von Wikipedia

Für eine automatische Kategorisierung von Texten auf Basis von Wikipedia identifiziert das *Open Topic Model* nach Waltinger und Mehler (Waltinger & Mehler, 2009) in einem ersten Schritt die Wikipedia-Artikel, die dem betrachteten Text am ähnlichsten sind, und extrahiert anschließend die diesen Artikeln zugeordneten Wikipedia-Kategorien. Anhand der taxonomischen Struktur des Kategoriensystems gelingt nun die Festlegung einer hierarchischen Reihenfolge der Kategorien und damit verbunden eine Generalisierung des Textthemas auf unterschiedlichen Ebenen. Einem Text, der sich mit der Person „Dirk Nowitzki“ beschäftigt, kann auf diese Weise mit der weiter unten in der Taxonomie angeordneten Kategorie „Basketballnationalspieler (Deutschland)“ ein sehr spezielles Thema zugeordnet werden, aber auch generellere Themen wie „Person (Würzburg)“ oder „Geboren 1978“.

Wie von Schönhofen (2006) gezeigt, kann eine Kategorisierung von Textdokumenten auch allein auf Grundlage der Titel und Kategorien von Wikipedia-Artikeln gelingen. Ermöglicht wird dies durch eine Bewertung aller Wikipedia-Kategorien anhand eines Vergleichs der Titel der Artikel, die den jeweiligen Kategorien zugeordnet sind, mit den im jeweiligen Dokument enthaltenen Wörtern. Beide bisher genannten Verfahren führen eine Kategorisierung auf Grundlage gesamter Texte durch. Sie liefern keinen Ansatz für einen Vergleich mehrerer Texte oder Textpassagen, der für eine automatische Bestimmung des aktuellen Dialogthemas jedoch erforderlich ist. Einen Lösungsansatz hierzu liefert die nachfolgend präsentierte Arbeit.

Ein von Chahine und Kollegen (2011) ebenfalls auf dem Kategoriensystem der Wikipedia basierender, weitaus komplexerer Ansatz unterstützt Bibliothekare bei der thematischen Einordnung pädagogischer Dokumente. In einem ersten Schritt wird für jedes Dokument ein Graph aufgebaut. Dafür werden alle in dem jeweiligen Dokument enthaltenen Wörter durch einen Termvergleich auf Konzepte in Form von Wikipedia-Artikeln abgebildet und die Kategorien und Oberkategorien dieser Artikel bis hin zum Wurzelknoten des Kategoriensystems extrahiert. Durch das Zusammenführen aller Termgraphen lässt sich das gesamte Dokument durch einen gerichteten, azyklischen Graphen darstellen. Wie in Abbildung 3.7 beispielhaft dargestellt, verlinken die einzel-

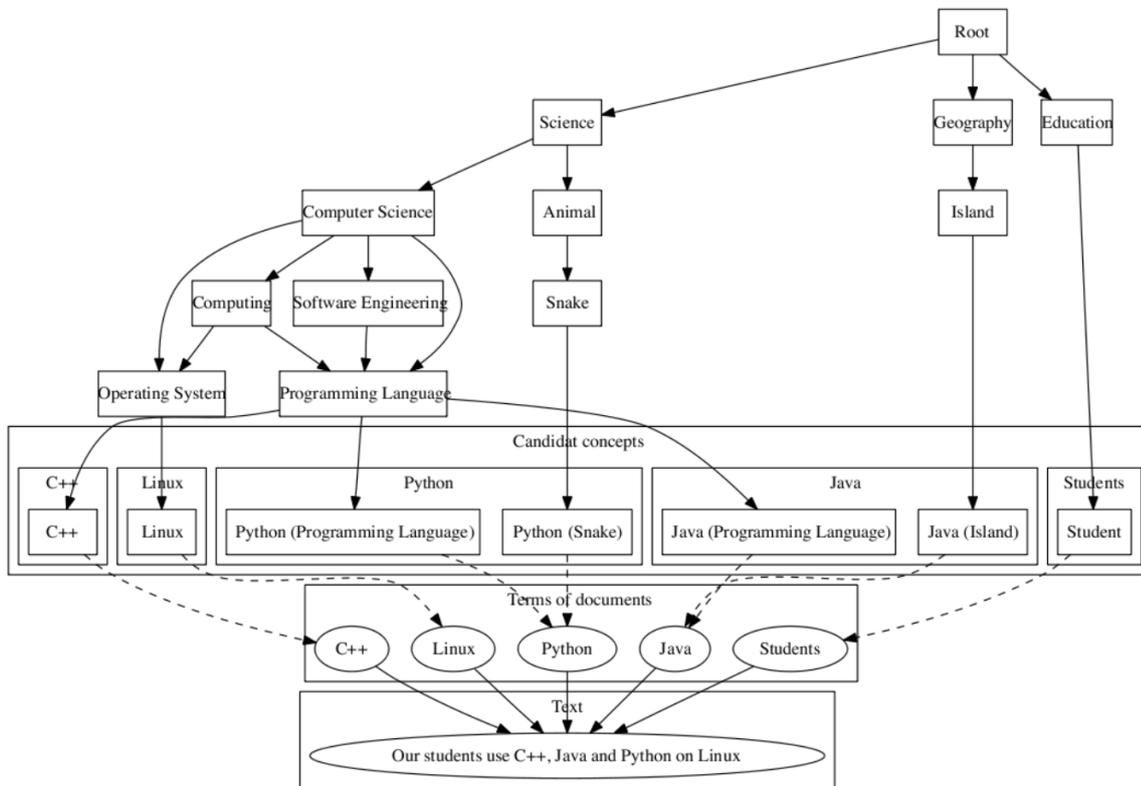


Abbildung 3.7: Wikipedia-basierte, graphische Darstellung des Satzes „Our students use C++, Java and Python on Linux.“ (Chahine et al., 2011, S. 197).

nen Terme des Dokuments auf die entsprechenden Wikipedia-Artikel, diese verlinken auf ihre Kategorien, letztere auf ihre Oberkategorien, etc., bis der Wurzelknoten der Wikipedia-Taxonomie erreicht ist. Das Thema des Dokuments lässt sich nun durch Spezifikation der speziellsten Kategorie, welche die meisten *bedeutungsvollen* Artikel subsumiert, bestimmen, wobei die Artikel als bedeutungsvoll definiert werden, die direkt oder indirekt mit einem Großteil der Satzterme verbunden sind. Auf diese Weise werden für das in Abbildung 3.7 dargestellte Beispiel die Themen „Computing“ und „Computer Science“ berechnet und gleichzeitig die mehrdeutigen Begriffe „Java“ und „Python“ entsprechend aufgelöst.

Der Vorteil dieses Ansatzes liegt, wie bei der von Ponzetto (2010) vorgestellten Herangehensweise zur Ähnlichkeitsbestimmung (s. Abschnitt 3.2.1), in der durch eine Abbildung der einzelnen Wörter auf Wikipedia-Artikel erreichte Konzeptualisierung und der Themenbestimmung auf Basis dieser Artikel. Zudem können Mehrdeutig-

keiten anhand des erkannten Themas aufgelöst werden. Da die von der Wikipedia bereitgestellte Menge an Kategorien, und somit die Menge möglicher Themen, allerdings sehr umfassend ist und mitunter sehr spezifische Titel wie „Kreisfreie Stadt in Nordrhein-Westfalen“ (in der deutschen Version) enthält, ist für eine Bestimmung des Dialogthemas eine vorausgehende Eingrenzung dieser Themenmenge erforderlich. Darüber hinaus ist dieses Verfahren für Texte zu einem Sachverhalt vorgesehen. Wie in Kapitel 2 ausführlich dargestellt, sind Dialoge jedoch geprägt durch häufige Themenwechsel, die sogar durch kohärente Beiträge herbeigeführt werden können. Das heißt, ein einzelner Beitrag kann Konzepte zum bisherigen und Konzepte für potentielle neue Themen enthalten, sodass häufig Wörter zusammen auftreten, die thematisch nicht miteinander zusammenhängen. Der hier vorgestellte Ansatz sucht für alle enthaltenen Konzepte eine einzige gemeinsame Kategorie, sodass eine Erkennung von sogenannten Themendrifts (s. Abschnitt 2.2.2) auf Grundlage dieses Verfahrens nicht realisierbar ist. Dieser Aspekt wird innerhalb der vorliegenden Arbeit durch eine Betrachtung der in den Dialogbeiträgen angesprochenen Konzepte als mögliche *Themenvorschläge* berücksichtigt.

Auch das von Bergstrom und Karahalios (2009) vorgestellte Verfahren zur Visualisierung thematisierter Sachverhalte in Besprechungen durch Abbildung der einzelnen Beiträge auf Wikipedia-Artikel vernachlässigt den innerhalb von Konversationen oft vollzogenen kohärenten Übergang zwischen den Themen. Allerdings zeigt diese Vorgehensweise, dass eine echtzeitfähige Themenerkennung auf Basis Wikipedias gelingen kann, indem das System die Informationen bezüglich des aktuellen konversationalen Themas mit jedem neuen Beitrag aktualisiert. Dies erfordert einen schnellen Zugriff auf das in Wikipedia enthaltene Wissen, der durch die in Abschnitt 3.1.4 vorgestellte Indexierung der benötigten Informationen gelingen kann.

3.2.3 Das TDT-Projekt

Das von der DARPA¹¹ finanzierte *Topic Detection and Tracking* (TDT)-Projekt (Allan, 2002) ist eine Initiative zur Entwicklung maschineller Verarbeitungsstrukturen, die die automatische Extraktion thematischer Strukturen aus unsegmentierten Sprachdatenströmen ermöglichen. Auf diese Weise lassen sich beispielsweise Radionachrichten ana-

¹¹*Defense Advanced Research Projects Agency*; Behörde des Verteidigungsministeriums der Vereinigten Staaten

lysieren und in einzelne, thematische Meldungen unterteilen. Die Daten liegen dabei in (transkribierter) Textform vor, obschon das langfristige Ziel des Projekts die direkte Bearbeitung von Audiodaten ist. Der gesamte Prozess der Themenerkennung und -verfolgung unterteilt sich in fünf Basisaufgaben, die im Weiteren genauer erläutert werden.

- **Segmentierung** Ziel dieser Aufgabe ist das Auffinden der Grenzen zwischen den einzelnen Meldungen und die damit einhergehende Unterteilung des Texts in seine einzelnen Themensegmente. Alle weiteren Aufgaben bauen auf dem Ergebnis der *Segmentierung* (engl. *story segmentation*) auf, da sie als Eingabe einzelne Segmente erwarten.
- **Themenentdeckung** Die Aufgabe der *Themenentdeckung* (engl. *first story detection*) ist die Ermittlung eines neuen, bisher noch nicht behandelten Nachrichtenthemas. Dies geschieht üblicherweise durch einen Vergleich der aktuellen Meldung mit den vorangegangenen anhand zuvor festgelegter Merkmale (beispielsweise durch einen Vergleich der Vektoren, s. Abschnitt 3.2.1).
- **Cluster-Erkennung** In Ergänzung zu der Themenentdeckung fasst dieser Prozess thematisch übereinstimmende Meldungen in einer Klasse zusammen. Wurde im Rahmen der Themenentdeckung ein neues Thema erkannt, besteht die Aufgabe der *Cluster-Erkennung* (engl. *cluster detection*) in der Definition einer neuen Klasse, wobei dem System im Vorfeld weder Anzahl noch Inhalte der festzulegenden Klassen bekannt sind.
- **Themenverfolgung** Die *Themenverfolgung* (engl. *tracking*) behandelt das Auffinden weiterer Meldungen zu einer bereits bestehenden Themenklasse. Auch die für diese Aufgabe bisher entwickelten Verfahren basieren typischerweise auf Merkmalsvergleichen.
- **Zusammenhangsermittlung** Zweck der *Zusammenhangsermittlung* (engl. *story link detection*) ist die Untersuchung, ob zwei willkürlich gewählte Meldungen dasselbe Thema besitzen. Diese Aufgabe bildet somit die Grundlage für alle weiteren Aufgaben und ist Bestandteil der Themenentdeckung und der Themenverfolgung.

Die einzelnen TDT-Aufgaben bilden einzelne, zum Teil in dieser Arbeit bereits eingeführte Forschungsfelder. Das im nachfolgenden Kapitel konzipierte Themenmodell führt Umsetzungen dieser Aufgaben zusammen mit dem Ziel, den sich in natürlich-sprachlichen Dialogen ereignenden thematischen Ablauf automatisch erfassen und spezifizieren zu können. Details über die genaue Definition der Aufgaben hinsichtlich der Zielanwendung sind Abschnitt 4.3.3 zu entnehmen.

3.3 Themenbewusstsein in der konversationalen Mensch-Maschine-Interaktion

Wie im vorherigen Abschnitt dargestellt, unterstützt eine maschinelle Themenerkennung beispielsweise die automatische Suche nach bestimmten Informationen durch eine thematische Einschränkung des gegebenen Suchraums oder die rechnergestützte Klassifikation von Textdokumenten durch die Bestimmung von Termähnlichkeiten. Durch eine Erweiterung der grundlegenden Erkennungsprozesse und der zusätzlichen Integration des daraus resultierenden Themenmodells in ein konversationales System kann sogar eine Emulation menschlichen Themenbewusstseins gelingen, die eine wie in 2.2.2 beschriebene thematische und dadurch kohärente Gesprächsführung zwischen einem Menschen und einer Maschine ermöglicht. Doch auch wenn die Geschichte der Entwicklung konversationaler Systeme bis in die Mitte der Siebziger Jahre zurückreicht (Cassell et al., 2000), mangelt es den meisten dieser sprachbasierten Schnittstellen noch heute an einem künstlichen Bewusstsein für einen adäquaten Umgang mit Themen und möglichen Themenwechseln (Glina & Kang, 2010). Dieser Abschnitt führt konversationale Systeme als natürliche, an den Menschen angepasste Benutzerschnittstellen ein und verdeutlicht die Relevanz eines künstlichen Themenbewusstseins für diese Systeme anhand bestehender Ansätze und Realisierungen.

3.3.1 Konversationale Schnittstellen

Die Konversation ist die natürlichste Art der menschlichen Kommunikation. Entsprechend hoch ist die Motivation, konversationale Schnittstellen zu schaffen, die eine dem Menschen angepasste, intuitive Interaktion mit Computern auf Basis natürlicher Sprache gewährleisten. Im Gegensatz zu anderen sprachbasierten Systemen zielen kon-

versationale Schnittstellen dabei auf einen dynamischen Dialog und einer informellen, sozialen Aktivität zwischen Mensch und Maschine ab (Oviatt & Adams, 2000). Nishida (2007) unterscheidet diesbezüglich zwischen einfachen Text-basierten Dialogsystemen, komplexeren konversationalen Agenten und hochentwickelten konversationalen Robotern. Eine der bekanntesten Dialogsysteme ist *ELIZA* (Weizenbaum, 1966), eine frühe Implementation einer einfachen natürlichsprachlichen Schnittstelle. Sowohl die Analyse der Benutzereingaben als auch die Generierung entsprechender Systemantworten erfolgen anhand von Regeln, die durch in der Eingabe enthaltene Schlüsselwörter selektiert werden. Konversationale Agenten sind künstliche Charaktere, die Dialogsysteme um einen virtuellen Körper und auf diese Weise um nonverbale Kommunikation erweitern (Nishida, 2007). Die natürlichste Interaktion zwischen Mensch und Maschine versprechen konversationale Roboter, die in der jeweiligen Umwelt verkörpert sind und daher einen gemeinsamen, physikalischen Interaktionsraum mit dem menschlichen Gegenüber aufbauen können (erste Ansätze hierzu liefern (Nguyen & Wachsmuth, 2011) und (Foster & Matheson, 2008)). Das Zielsystem des im Rahmen der vorliegenden Arbeit entwickelten Themenmodells ist ein konversationaler Agent, sodass diese Form einer konversationalen Schnittstelle und die Motivation hinter der Wahl dieses Systems als Plattform im Folgenden weiter spezifiziert werden.

Konversationale Agenten

Die Erkenntnis, dass Menschen konversationalen Schnittstellen weniger als Benutzer, sondern eher als soziale Partner entgegentreten (Nass, Steuer & Tauber, 1994), hat zu der Entwicklung animierter Charaktere beigetragen (Oviatt & Adams, 2000), die gemeinhin als *konversationale Agenten* (im Sinne von *Embodied Conversational Agents* nach (Cassell et al., 2000)) bezeichnet werden. Wie in Abschnitt 1.1 dargestellt, sind diese Agenten virtuelle Schnittstellen, deren Modalitäten mit denen der menschlichen Konversation übereinstimmen. Folglich sind sie dazu in der Lage, sowohl verbale als auch nonverbale Äußerungen zu verstehen und zu generieren und Dialogfunktionen wie Sprecherwechsel (engl. *turn-taking*) und Rückmeldungen (engl. *feedback*) zu erfüllen (Cassell et al., 2000). Dies ermöglicht den menschlichen Anwendern, auf natürliche, intuitive Art mit der Maschine zu interagieren, nämlich in einem direkten persönlichen Dialog. Auf diese Weise wird nachweislich das menschliche Wohlbefinden und die Aufgeschlossenheit gegenüber Computersystemen erhöht (Foster, 2007) sowie eine

Verbesserung der Benutzerfreundlichkeit von Mensch-Maschine-Schnittstellen erreicht. Nachfolgend werden exemplarisch drei konversationale Agenten vorgestellt. Eine umfangreichere Übersicht ist in (Kopp, 2003) zu finden.

Die konversationale Agentin **Greta** (De Carolis, Rosis, Carofiglio, Pelachaud & Poggi, 2001) wird durch einen sogenannten *Talking Head*, das heißt durch einen virtuell dargestellten, menschlichen Kopf verkörpert (s. Abbildung 3.8a¹²), und versorgt ihren menschlichen Dialogpartner mit Informationen hinsichtlich einer gegebenen Domäne. Zu Beginn einer Interaktion wird je nach Situation und Kontext ein dem gesamten Dialog übergeordnetes Ziel gesetzt, beispielsweise „Informiere deinen Dialogpartner über Therapiemöglichkeiten bei Angina!“. Dieses Ziel gibt das Thema des Dialogs vor und aktiviert bestimmte, dem jeweiligen Thema zugeordnete Konversationspläne, auf dessen Grundlage das System Fragen des jeweiligen Benutzers zu diesem Thema beantworten kann. Dabei zeigt Greta je nach Dialogthema und Verhältnis zu ihrem aktuellen menschlichen Gegenüber Emotionen, die durch entsprechende Mimiken zum Ausdruck kommen.

Rea (*Real Estate Agent*) (Cassell et al., 1999), dessen Realisierung den Begriff *Embodied Conversational Agent* prägte, ist eine am MIT entwickelte, virtuelle Immobilienmarklerin (s. Abbildung 3.8b¹³). Die Agentin wird auf einem großen Bildschirm angezeigt, vor dem der menschliche Anwender durch Kamera- und Mikrofoneingaben mit dem System interagiert. Durch Zugriff auf eine spezielle Datenbank kann Rea Bilder von Immobilien und ihren einzelnen Räumen anzeigen und über spezielle Merkmale diskutieren. Dabei werden sowohl Verständnis als auch Generierung von natürlichsprachlichen Gesprächsbeiträgen durch Regeln ermöglicht und weitere Dialogfunktionen auf sogenannte Diskursfunktionen (engl. *discourse functions*) abgebildet, die während des Dialogs ausgeführt werden.

Die beiden vorgestellten Agentinnen sind auf aufgabenabhängige Konversationen mit ihren menschlichen Gesprächspartnern ausgerichtet und die jeweiligen möglichen Dialogverläufe somit hinsichtlich der gegebenen Aufgabe thematisch eingeschränkt. Agentensysteme dieser Art handhaben Benutzereingaben, die außerhalb ihres Wissensbereichs liegen, indem sie ihre Unfähigkeit hinsichtlich der Verarbeitung dieser Äußerung zum Ausdruck bringen oder diese ganz einfach missachten (Mehta & Cor-

¹²Quelle: <http://perso.telecom-paristech.fr/~pelachau/Greta/>

¹³Quelle: <http://www.media.mit.edu/gnl/projects/humanoid/>

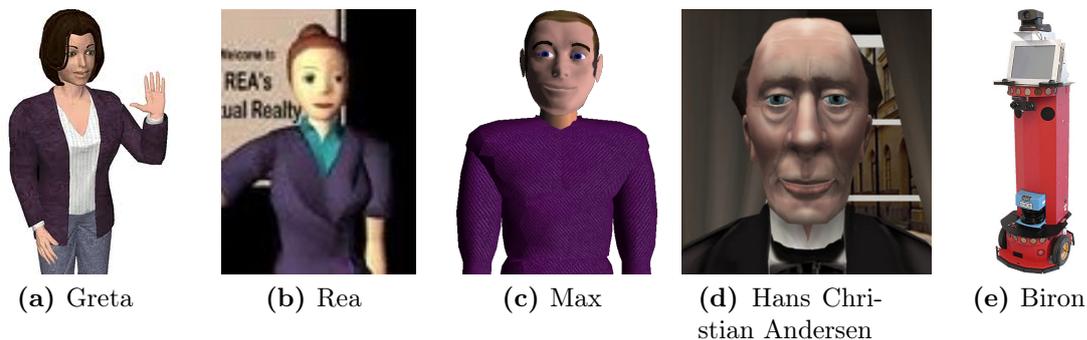


Abbildung 3.8: Beispiele konversationaler Agenten.

radini, 2008). Ein gänzlich anderer Ansatz ist die Anbindung des Agenten an eine externe Informationsquelle, die den systeminternen Wissensbestand erweitert. Ein Beispiel hierfür liefert der in der Arbeitsgruppe *Wissensbasierte Systeme* der Universität Bielefeld entwickelte, multimodale Agent **Max** (s. Abbildung 3.8c¹⁴). Die Dialogfähigkeit des unter anderem als Museumsführer eingesetzten Agenten wird durch manuell erstelltes Dialogwissen realisiert, welches aus Regeln für die Interpretation der natürlichsprachlichen Eingaben und für die darauf aufbauende Generierung geeigneter, multimodaler Ausgaben besteht. Die Auswahl der Regeln geschieht durch einen Musterabgleich (engl. *pattern matching*) der in der Eingabe enthaltenen Wörter und den in den einzelnen Regeln definierten Mustern. Mit dem Ziel, thematisch uneingeschränkte, flexiblere Dialoge zwischen Max und seinem menschlichen Gegenüber zu ermöglichen, beinhaltet das Konversationswissen des Agenten zusätzliche Smalltalk-Regeln, die das für das gegebene Museumsszenario typische Domänenwissen um Alltagsthemen erweitert (Kopp et al., 2005). Darüber hinaus erlaubt die während der Interpretation durchgeführte Klassifikation der Eingabe (z. B. in Begrüßung, Frage, Verneinung, etc.) eine adäquate Antwortgenerierung, auch wenn eine vollständige Interpretation der vom Menschen eingegebenen Äußerung fehlschlägt. So antwortet Max auf eine unvollständig interpretierte Frage beispielsweise mit „*Das ist aber eine interessante Frage!*“.

Eine im Zuge des KnowCIT-Projekts (s. Abschnitt 1.3) erfolgte Anbindung von Max an Wikipedia und dessen von der DBpedia (s. Abschnitt 3.1.3) bereitgestellten ontologischen Darstellung erweitert das manuell erzeugte Dialogwissen des Agenten um lexikalische Inhalte (Waltinger et al., 2012). Dies gelingt über eine Question Ans-

¹⁴Quelle: Archiv Arbeitsgruppe Wissensbasierte Systeme, Universität Bielefeld

wering (QA)-Schnittstelle, die durch ein Schlüsselwort zu Beginn einer Eingabe angesprochen wird. Das heißt, wenn der menschliche Dialogpartner eine offene Frage an das System stellt und diese mit „wikiqa:“ (*WikiQA* ist der Name des QA-Systems¹⁵) einleitet, wird diese mittels einer entsprechenden Regel an das QA-System weitergeleitet. Letzteres überträgt die in natürliche Sprache formulierte Frage zunächst in eine RDF-Tripeldarstellung (s. Abschnitt 3.1.3) und bildet diese auf die in DBpedia enthaltenen Strukturen ab, um die gewünschte Information abzufragen. Kann keine Übereinstimmung festgestellt werden, wird in den Texten der Wikipedia-Artikel nach der passenden Information gesucht. Zu diesem Zweck wird die eingegebene Frage mit den über drei Millionen in den Wikipedia-Artikeln enthaltenen Sätzen abgeglichen. Der Satz mit der höchsten Übereinstimmung hinsichtlich der enthaltenen Terme wird von dem Agenten als Antwort vorgetragen. Dem Agentensystem steht auf diese Weise eine zusätzliche Informationsquelle mit lexikalischen Inhalten zur Verfügung, deren Aktualität, Erweiterung und Optimierung durch freiwillige Wikipedia-Autoren gewährleistet werden. Eine ausführlichere Darstellung des an den Agenten Max angebundenen QA-Systems liefert (Waltinger et al., 2012).

3.3.2 Beispielsysteme für künstliches Themenbewusstsein

Der Dialog zwischen Mensch und Maschine wird in konversationalen Schnittstellen im Allgemeinen durch Pläne ermöglicht, die Regeln für die Verarbeitung natürlichsprachlicher Eingaben und Regeln für die Generierung geeigneter Ausgaben beinhalten (wie auch bei dem vorab beschriebenen Agenten Max). Auf diese Weise lassen sich aufgabenorientierte Dialoge als Pläne modellieren, die thematisch annotiert werden können. So gelang schon früh eine Spezifikation der in den Dialogen angesprochenen Themen durch Beobachtung der Abfolge der in dem jeweils betrachteten Dialog aufgerufenen Pläne (Litman & Allen, 1987). Die bisherige Themenbestimmung des Agenten Max setzt auf diesen Ansatz auf und ermöglicht zudem eine simple Umsetzung eines künstlichen Themenbewusstseins des Agenten durch eine themenbedingte Generierung der Agentenäußerungen.

Anstelle einer thematischen Annotation der Dialogpläne, überführt der von Glinn und Kang (2010) vorgestellte *Ripple Down Rules* (RDR)-Algorithmus die einzelnen Dialogregeln in eine Baumstruktur, welche die thematischen Zusammenhänge dieser

¹⁵s. <http://www.wikiqa.de>

Regeln widerspiegelt. Simultan zum aktuell laufenden Dialog wird auf Grundlage dieser Baumstruktur ein neuer Baum aufgebaut, dessen Knoten die aktivierten Regeln repräsentieren. Ausgehend vom zuletzt besuchten Knoten, sucht das System für die anschließende Eingabe zunächst bei den jeweiligen Unterknoten nach einer passenden Regel für die Verarbeitung dieser Eingabe. Auf diese Weise geht das System zunächst von einer Vertiefung des aktuellen Themas aus. Liefert keiner der Unterknoten eine passende Regel, interpretiert das System dies als Themenwechsel und durchsucht alle weiteren, bisher inaktiven Einträge. Eine Übereinstimmung nahe der Wurzel der gesamten Baumstruktur weist dabei auf ein neues Thema hin, ein Treffer weiter unten im Baum wird als Wiederaufnahme eines bereits besprochenen Themas interpretiert. So ermöglicht der RDR-Algorithmus dem System ein automatisiertes Bewusstsein für Kohärenz und mögliche Themenwechsel, inhaltliche und konzeptuelle Informationen bezüglich der Dialogthemen bleiben dem System dabei jedoch vorenthalten. Hinzu kommt das Problem jeder auf Basis von Regeln durchgeführten Themenerkennung: Sie funktioniert nur, solange die Themen nah an der Aufgabenstruktur liegen, beziehungsweise die einzelnen Äußerungen auf thematisch annotierte oder sortierte Regeln abgebildet werden können (Litman & Allen, 1987). Ein großer Vorteil des Algorithmus liegt in der Auflösung mehrdeutiger Verweise, sodass das jeweilige konversationale System Nachfragen der Art „*Wie funktioniert das?*“ oder Aufforderungen wie „*Erzähl mir mehr!*“ auflösen und beantworten kann.

Einen ganz anderen Ansatz verfolgen Wang und Jiang (2003). Sie gruppieren und organisieren die Wörter, die das betrachtete konversationale System akzeptiert, in einer Taxonomie (s. Abschnitt 3.1.1). Wörter, die in dieser Hierarchie weiter oben angeordnet und somit genereller sind, werden als Konzeptwörter definiert. Zusätzlich greift das System auf eine vorab definierte Baumstruktur an möglichen Themen zu, bei der jedem Themenknoten eine Menge an Konzeptwörtern zugeordnet ist. Eine Benutzeräußerung wird thematisch bestimmt, indem die in der Eingabe enthaltenen Wörter auf Konzeptwörter der Themenknoten abgebildet werden. Unter der Annahme, dass sich die Dialogthemen entlang des Baums entwickeln, formen die im Dialog genannten Themen einen Unterbaum. Nachfolgende Dialogthemen werden erneut auf Grundlage der enthaltenen Wörter erkannt, zusätzlich jedoch mit möglichen Kandidaten, das heißt Unterknoten, die die Baumstruktur vorgibt, abgeglichen. Der Knoten mit der höchsten Übereinstimmung repräsentiert das aktuelle Thema. Diese Form der

Themenbestimmung liefert mit 83,8% Korrektheit sehr zuverlässige Ergebnisse und wie beim RDR-Algorithmus gelingt auch hier mit Hilfe der Baumstruktur eine Prognose hinsichtlich der Themen, die der jeweilige Benutzer mit der nächsten Äußerung ansprechen könnte. Die Vorteile dieser Methode können jedoch aufgrund des hohen Aufwands für den manuellen Aufbau des Themenbaums und der Taxonomie nur für eine eingeschränkte Dialogdomäne gewährleistet werden. Eine Emulation künstlichen Themenbewusstseins in natürlichsprachlichen Smalltalk-Situationen kann somit mit keinem der beiden vorgestellten Ansätzen gelingen, da diese den Fokus auf aufgabenorientierte Dialoge legen und nur einen geringen Spielraum bei der Wahl der Wörter und der Dialogthemen zulassen. Der nachfolgende Abschnitt befasst sich daher mit Ansätzen automatischen Themenbewusstseins für konversationale Agenten mit Smalltalk-Funktionalitäten.

Themenbewusste konversationale Agenten

Menschen verfügen über konversationale Fähigkeiten, die die für eine erfolgreiche Interaktion zwischen zwei oder mehr Personen erforderliche Zusammenarbeit der beteiligten Interaktionspartner ermöglichen. Ein konversationaler Agent sollte bestenfalls dazu in der Lage sein, diese Fähigkeiten zu emulieren und dem Benutzer ein Gefühl der Kooperation während des gemeinsamen Dialogs zu geben (Novielli & Strapparava, 2011). Dazu gehört auch, mit dem menschlichen Gegenüber eine thematische, kohärente Unterhaltung aufzubauen und über die von der gegebenen Dialogsituation abhängigen Angemessenheit von Themen zu entscheiden. Das System der zuvor beschriebenen Agentin Rea beispielsweise präferiert Systemausgaben, die das aktuelle Thema aufrecht halten und wählt neue Themen abhängig von dem Grad der Freundschaft zwischen der Agentin und ihrem menschlichen Dialogpartner aus (Cassell et al., 1999)¹⁶. Nach Patel, Leuski und Traum (2006) erhöht ein virtueller Charakter sogar die Glaubwürdigkeit seiner Teilnahme an dem Dialog, wenn er sogenannte *off-topic* Äußerungen seines Gegenübers, also Äußerungen, die nicht dem aktuellen Dialogthema entsprechen, erkennt und dies durch entsprechende Reaktionen verdeutlicht. Diese Art der konversationalen Fähigkeiten eines Agentensystems setzt ein künstliches Bewusstsein für Themen in Dialogen voraus, die über die rein regelbasierte Themenerkennung, wie sie beispielsweise bei Rea implementiert ist, hinausgeht. Wie bereits

¹⁶Die genaue Umsetzung bleibt dem Leser leider vorenthalten.

erwähnt, existieren derzeit nur sehr wenige (dokumentierte) Ansätze zur Simulation menschenähnlichen Themenbewusstseins in konversationalen Agenten, die im Folgenden vorgestellt werden.

Eine im Rahmen des europäischen *NICE*¹⁷-Projekts entwickelte 3D-animierte Darstellung des Schriftstellers **Hans Christian Andersen** (s. Abbildung 3.8d¹⁸) ermöglicht Kindern und Jugendlichen eine natürliche, multimodale Interaktion mit einem virtuellen Märchenautor und seinen Märchenfiguren (Bernsen & Dybkjær, 2005). Ziel dieser Anwendung ist die Vermittlung von Informationen über das Leben, das Zeitalter und die Märchen des Autors. Hierfür wird anhand von Regeln zunächst die Funktion der betrachteten Eingabe (Frage, Kommentar, etc.) ermittelt und den verbleibenden Wörtern anschließend mittels Musterabgleich zuvor manuell erstellten Kategorien zugeordnet. Auf Grundlage dieser Information und einer Menge an weiteren Regeln gelingt die Auswahl einer passenden Antwort für den virtuellen Agenten. Um den Unterhaltungsfaktor des Agenten zu erhöhen, implementierten Mehta und Corradini (2008) einen Webagenten, welcher mit Hilfe bestehender QA-Systeme die Beantwortung von Fragen gewährleistet, die von der eigentlichen Domäne des konversationalen Systems abweichen und für die daher keine passenden Antwortregeln in der Wissensbasis verfügbar sind. Darüber hinaus ermöglicht ein maschineller Zugriff auf die von Google bereitgestellte Verzeichnisstruktur, die sogenannten *Google Groups*, eine thematische Einordnung der in den Eingaben enthaltenen Wörter in unterschiedliche Kategorien. Mit dieser zusätzlichen Themenbestimmung soll in Zukunft eine Basis für einen wechselseitigen Dialog zwischen dem künstlichen und dem menschlichen Dialogpartner über das jeweilige, vom Benutzer angesprochene Thema geschaffen werden. Vorstellbar wäre auch eine thematische Eingrenzung möglicher Antworten, die eine gezielte Auswahl des zu verwendenden QA-Systems und auf diese Weise eine Verringerung der Antwortzeit des Agenten erlaubt. Einen Ansatz hierzu stellen Waltinger, Breuing und Wachsmuth (2012) für den bereits eingeführten Agenten Max vor. Die Autoren grenzen die Menge möglicher Wikipedia-Artikel, die bei einer Frage an das QA-System des Agenten nach passenden Antworten durchsucht werden, thematisch ein, indem nur solche Artikel bei der Suche betrachtet werden, die einer bestimmten Kategorie zugeordnet sind. Die Bestimmung der Kategorie (und somit des Themas) der Benutzereingabe erfolgt

¹⁷Natural Interactive Communication for Edutainment

¹⁸Quelle: http://www.multimodalusability.dk/mus_support_andersen_system.php

dabei mit Hilfe des zur Kategorisierung von Texten eingesetzten *Open Topic Models* (Waltinger & Mehler, 2009) (s. Abschnitt 3.2.2).

Ein von Maas (2007) eingeführter, komplett anderer Ansatz zielt darauf ab, Objektreferenzen auf Grundlage automatischer Themenerkennung besser auflösen zu können und das betrachtete System auf diese Weise mit einem Bewusstsein für unterschiedliche Situationen auszustatten. Bei dem Zielsystem handelt es sich um den mobilen Roboter **Biron** (Bielefeld Robot Companion (Haasch et al., 2004), s. Abbildung 3.8e¹⁹), eine Plattform zur Erforschung natürlicher, aufgabenbezogener Interaktion zwischen Menschen und Service-Robotern. Informationen über das aktuelle Dialogthema unterstützen einen als Haushaltshilfe eingesetzten Roboter beispielsweise bei der Entscheidung, ob mit dem Begriff „Bürste“ eine Haar- oder eine Spülbürste gemeint ist. Voraussetzung für eine derartige Auflösung mehrdeutiger Begriffe liefert ein Themenerkennungssystem, dessen Prozesse die in dem TDT-Projekt definierten Aufgaben für automatische Themenerkennung (s. Abschnitt 3.2.3) ausführen. Die Ergebnisse werden zur Laufzeit (*online*) und die Themen dabei *dynamisch* generiert, um sich wechselnden Umgebungen anzupassen. Dynamisch bedeutet hier, dass der Erkennungsprozess nicht auf eine Menge an vordefinierten Themen zugreift, sondern die Menge möglicher Themen auf Basis des laufenden Dialogs aufgebaut wird. Für eine Bestimmung des aktuellen Themas werden hier die sprachlichen Beiträge des menschlichen Interaktionspartners verarbeitet, weitere Informationen liefert der Objekterkenner des Roboters. Welches Merkmal (Wort oder Objekt) zu welchem Thema gehört wird anhand von *Kookkurrenzen*, also dem gemeinsamen Auftreten von zwei Merkmalen, festgelegt. Dabei werden Merkmale, die oft zusammen im selben Kontext auftauchen, als zu einem Thema gehörend erachtet und umgekehrt, was eine dynamische Erstellung von Themen durch eine Einteilung der Merkmale in unterschiedliche, vorher unbekannte Gruppen (engl. *cluster*) ermöglicht. Die *Cluster-Erkennung* (s. Abschnitt 3.2.3) erfolgt dabei global, das heißt die erforderlichen Prozesse beziehen die gesamte Datenbasis (hier: alle bisherigen Äußerungen) in die Analysen mit ein, statt, wie bei den lokalen Verfahren, einen Vergleich der neu eintreffenden Datenstruktur (Äußerung) mit bereits existierenden Clustern vorzunehmen. Dies geht auf Kosten der Echtzeitfähigkeit des Systems, Grenzen (Wechsel) zwischen den einzelnen Themenblöcken lassen sich jedoch zuverlässiger ermitteln. Zudem ist keine Identifikation kohärenter Themenwechsel möglich, da keine

¹⁹Quelle: <http://www.ubergizmo.com/2006/11/i-robot-you-human/>

Suche nach Themenüberlappungen auf Konzeptebene vorgenommen wird.

Künstliches Themenbewusstsein ist auch Gegenstand der Arbeit von Endrass, Rehm und André (2011), bei der eine automatische Anpassung des Kommunikationsverhaltens konversationaler Agenten an die kulturellen Hintergründe ihrer menschlichen Dialogpartner angestrebt wird. Hierfür wurden die in einer Studie zur Analyse kultureller Verhaltensunterschiede zwischen Deutschen und Japanern in typischen Smalltalk-Situationen zusammengetragenen Ergebnisse auf virtuelle Agenten übertragen. Zu den beobachteten und implementierten Unterschieden zählten neben Erscheinungsbild und nonverbalen Handlungsweisen wie Gesten und Körperhaltung insbesondere die Wahl der Gesprächsthemen, wie in Abschnitt 2.2.2 bereits dargestellt. Demnach adressierten die deutschen Probanden vorrangig allgemeine Themen, wie das Wetter oder die aktuellen Sportereignisse, aber auch persönliche Themen, wie Hobbies oder Berufe der Gesprächsteilnehmer. Die Japaner hingegen legten den Fokus auf allgemeine Themen und beschäftigten sich in ihren Dialogen nur sehr wenig mit privaten Inhalten. Ein weiterer Unterschied ließ sich hinsichtlich der Abarbeitung der einzelnen besprochenen Themen und der Häufigkeit von Themenwechseln feststellen. So konnten bei beiden Kulturen insgesamt 26 unterschiedliche Themen identifiziert werden, die bei den Deutschen von 38 und bei den Japanern von 46 Themenwechseln begleitet wurden. Dieses Ergebnis entspricht der von Ting-Toomey im Jahre 1999 aufgezeigten Unterscheidung in *monochrome* und *polychrone* Kulturen und somit der davon abgeleiteten Tendenz, dass Japaner häufiger zwischen den Themen hin- und herspringen, während westliche Personen Gesprächsinhalte nacheinander abarbeiten und weniger Themenwechsel durchführen (Endrass et al., 2011). Eine Pilotstudie zur Prüfung, ob die genannten Erkenntnisse hinsichtlich der kulturspezifischen Unterschiede auch auf virtuelle Agenten zutreffen, ließ Differenzen in der Wahrnehmung der Dialoge mit Agenten, die die kulturspezifischen Verhaltensweisen des jeweiligen Probanden emulierten, mit solchen, bei dem der Agent Verhaltensweisen einer anderen Kultur aufwies, erkennen.

3.4 Zusammenfassung

Gegenstand dieses Kapitels war die Vorstellung elementarer Methoden und Ansätze für die technische Umsetzung künstlichen Themenbewusstseins und die Aufarbeitung des aktuellen Stands der Forschung auf diesem Gebiet. Zu Beginn wurde der für eine

automatische Themenerkennung erforderliche Zugriff auf die Repräsentation maschinenlesbaren Konzeptwissens behandelt und bestehende Quellen für den Aufbau einer solchen Wissensrepräsentation diskutiert. Dabei lieferten insbesondere die Vorteile der Aktualität und Objektivität die Gründe für die Wahl der Online-Enzyklopädie Wikipedia als geeignete Grundlage für den Aufbau eines maschinenlesbaren, mentalen Lexikons.

Im weiteren Verlauf des Kapitels wurden die Bestimmung von Textähnlichkeiten und die Textkategorisierung als klassische Methoden zur automatischen Erkennung von Themen in natürlichsprachlichen Texten besprochen. Der Fokus lag dabei auf der Beschreibung von Ansätzen, die diese Methoden anhand der Wissensstrukturen der Wikipedia realisieren. Ziel war hierbei die Erfassung von Ideen für die technische Umsetzung der automatischen Erkennung von Themen auf der Grundlage des Wikipedia-Kategoriengraphen. Den Abschluss dieses Abschnitts bildete die Einführung des TDT-Projekts und der Beschreibung der darin definierten Aufgaben zur automatischen Extraktion thematischer Strukturen aus unsegmentierten Sprachdatenströmen. Die hierbei vorgenommene Unterteilung des Erkennungsprozesses in unterschiedliche Aufgaben wird für den Prozess der Themenerkennung übernommen, wobei die einzelnen Aufgaben dabei der Identifikation der unterschiedlichen Themeninformationen dienen.

Abschließend wurden bestehende Ansätze zur Umsetzung künstlichen Themenbewusstseins in konversationalen Systemen betrachtet. In diesem Kontext wurden konversationale Agenten als natürliche Schnittstellen zur Interaktion mit Maschinen eingeführt und einzelne Ausführungen dieser Systeme, unter anderem der Agent Max, vorgestellt. Die daran anschließende Auswertung des aktuellen Stands der Forschung hinsichtlich themenbewusster Agenten verdeutlichte, dass die für eine Realisierung künstlichen Themenbewusstseins erforderliche Themenerkennung bislang auf Basis von Dialogregeln gelang. Die damit verbundenen Nachteile wie der hohe manuelle Aufwand und die Einschränkung auf bestimmte Domänen veranschaulichten die Notwendigkeit eines von den Regeln losgelösten Themenerkenners.

Im nächsten Kapitel folgt die Darstellung der Konzeption eines Themenerkenners, der dieser Forderung nachkommt und die für die Emulation menschenähnlichen Themenbewusstseins erforderlichen Informationen ermittelt. Durch die daran anschließende Integration dieser Informationen in das konversationale Verhalten des Agenten

Max gelingt schließlich die Realisierung künstlichen Themenbewusstseins in natürlichen Dialogen zwischen Mensch und Agent.

4 Künstliches Themenbewusstsein für den konversationalen Agenten Max

Nachdem in den letzten beiden Kapiteln die für die Konzeption eines Systems zur Emulation menschlichen Themenbewusstseins in künstlichen Agenten erforderlichen theoretischen und technischen Grundlagen vermittelt wurden, führt dieses Kapitel das der vorliegenden Arbeit zugrunde liegende Gesamtkonzept ein und liefert einen Überblick über konkrete Konzeptentscheidungen. Dementsprechend beginnt dieses Kapitel mit einer kurzen Wiederholung der Aufgabenstellung. Darauf aufbauend werden die Gesamtarchitektur und dessen Einteilung in zwei Aufgabenschwerpunkte für die technische Umsetzung der beschriebenen Aufgabe vorgestellt. Anschließend erfolgt für jeden der beiden Aufgabenschwerpunkte die Darstellung der Anforderungen, die sich aus der Aufgabenbeschreibung ergeben, die Einführung einer entsprechenden Architektur und eine Diskussion der jeweiligen Konzeptentscheidungen.

4.1 Aufgabenbeschreibung

Wie in Kapitel 1 aufgezeigt, ist das Ziel dieser Arbeit die Entwicklung eines Systems zur Emulation menschlichen Themenbewusstseins in natürlichen Dialogen. Voraussetzung dafür ist die automatische Erkennung von Informationen hinsichtlich des aktuellen Themas des laufenden Dialogs und die Integration dieser Informationen in das konversationale Verhalten des künstlichen Agenten Max. Die Themenerkennung ist dabei so umzusetzen, dass eine Identifikation der in Abschnitt 2.2.2 zusammengefassten, den thematischen Ablauf eines Dialogs beeinflussenden Aktionen gelingt. Das daraus gewonnene Wissen ist dann so in das Dialogsystem des Agenten zu integrieren, dass dessen kommunikativen Fähigkeiten durch dieses Bewusstsein erweitert und verbessert werden. Dadurch soll Max die thematischen Aktionen innerhalb eines Dialogs sowohl wahrnehmen als auch anhand adäquater Äußerungen initiieren können.

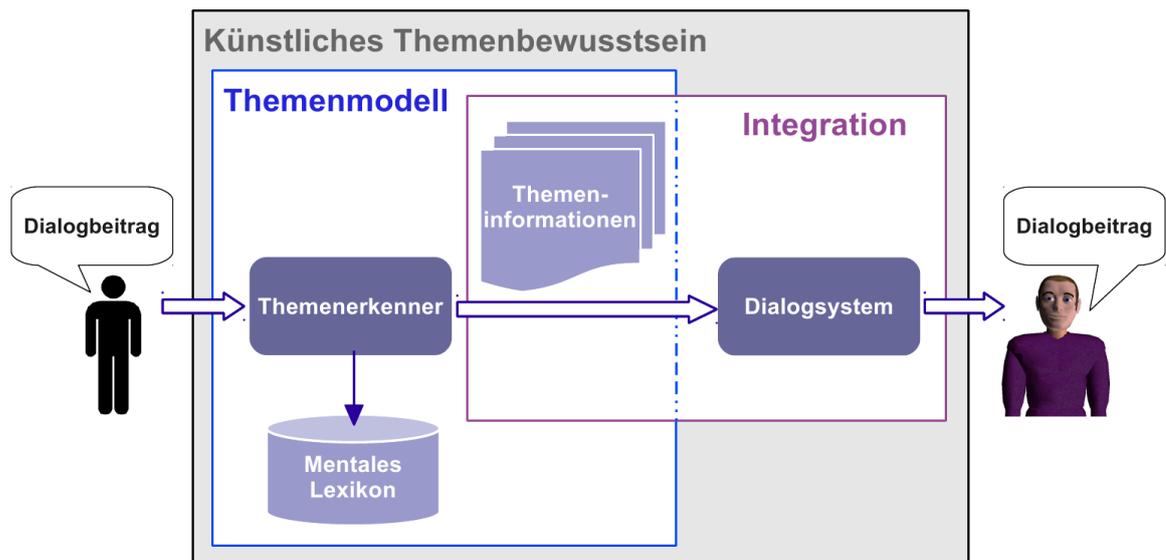


Abbildung 4.1: Gesamtarchitektur, auf dessen Grundlage die Realisierung künstlichen Themenbewusstseins für den konversationalen Agenten Max gelingt. Die Architektur besteht im Wesentlichen aus zwei Komponenten: dem *Themenerkennung* zur Identifikation und Generierung aktueller Themeninformationen und dem *Dialogsystem* des Agenten, in welches die erkannten Informationen integriert werden.

4.2 Das Gesamtkonzept

Wie bereits in Abschnitt 2.2.2 dargestellt, lässt sich Themenbewusstsein beschreiben als menschliche Fähigkeit, Themen sowohl wahrzunehmen als auch in das individuelle Gesprächsverhalten einzubeziehen. Die Umsetzung emulierten Themenbewusstseins für den Agenten Max setzt sich demnach aus zwei Herausforderungen zusammen: Erstens, die Erkennung und Analyse von Gesprächsthemen in natürlichen Dialogen, und zweitens, die Generierung thematischen Konversationsverhaltens des künstlichen Agenten. Auf Grundlage dieser Herausforderungen lassen sich zwei Aufgabenschwerpunkte formulieren:

1. **Konzeption und Umsetzung eines Themenmodells zur Generierung aktueller Themeninformationen**
2. **Integration der gewonnenen Themeninformationen in das Dialogsystem des Agenten Max**

Abbildung 4.1 liefert einen Überblick über die aus der Aufgabenstellung hervorgegangene Gesamtarchitektur sowie deren Unterteilung in zwei Aufgabenschwerpunkte.

Das **Themenmodell** besteht demnach aus einem automatischen Themenerkennner, der auf Basis natürlichsprachlicher Dialogbeiträge und unter Zugriff auf maschinenlesbares Konzeptwissen in Form eines künstlichen mentalen Lexikons aktuelle Themeninformationen ermittelt. Durch eine **Integration** dieser Informationen in das Dialogsystem des Agenten Max gelingt die Generierung adäquater Äußerungen des Agenten. Nachfolgend werden die Konzepte für die Umsetzung der beiden Aufgabenschwerpunkte eingeführt.

4.3 Das Themenmodell

Die Hauptkomponente des Themenmodells bildet der **Themenerkennner**, dessen Aufgabe in der Erkennung des aktuellen Dialogthemas auf Basis der einzelnen Dialogbeiträge besteht. Wie in Abschnitt 2.3 festgehalten, kann eine erfolgreiche Themenerkennung nur durch Zugriff auf grundlegendes Wissen über Konzepte und ihre Zusammenhänge gewährleistet werden. Eine weitere Komponente des Themenmodells bildet somit ein künstliches **mentales Lexikon**, welches maschinenlesbares Konzeptwissen bereitstellt. Im Zuge der automatischen Erkennung erfolgt die Ermittlung weiterer Informationen hinsichtlich der aktuellen Themensituation des betrachteten Dialogs. Diese **Themeninformationen** bilden die Ausgabe des Themenmodells und werden an das jeweilige Zielsystem weitergeleitet.

Ziel dieses Abschnitts ist die Darstellung des aus den theoretischen und technischen Grundlagen hervorgegangenen Konzepts für das zu entwickelnde Themenmodell. Grundlage dafür liefert die Erläuterung, was innerhalb der vorliegenden Arbeit unter dem Begriff „Thema“ innerhalb eines Dialogs verstanden wird. Demnach erfolgt zunächst die Einführung der Definition eines *Dialogthemas*. Im Anschluss daran wird ein Konzept für den automatischen Themenerkennner vorgestellt. Hierfür werden zunächst die an den Themenerkennner gestellten Anforderungen zusammengetragen und darauf aufbauend eine Architektur entwickelt. Danach folgt eine Beschreibung der Themeninformationen, die anhand des Themenerkennners ermittelt werden sollen. Anhand dieser erfolgt die Spezifikation von Aufgaben, die der Themenerkennner zusätzlich im Rahmen der Erkennungsprozesse ausführen muss.

4.3.1 Arbeitsdefinition „Dialogthema“

Auf Grundlage der in Abschnitt 2.2.1 erörterten Standpunkte bezüglich des Begriffs „Thema“ und angesichts der Tatsache, dass sich die im weiteren Verlauf der Arbeit betrachteten Diskurse auf Dialoge beschränken, lässt sich folgende konkrete Arbeitsdefinition eines *Dialogthemas* erschließen:

Definition Dialogthema

Ein **Dialogthema** ist eine unabhängige, selbstgewählte Kategorie, welche die Konzepte einer kooperativ erbrachten Sequenz an Dialogbeiträgen subsumiert.

Diese Definition wurde hauptsächlich von Bublitz (1989) übernommen, erweitert diese jedoch um den von Svennevig (1999) angefügten Aspekt der Zusammenfassung der angesprochenen Konzepte zu einer *subjektiv* gewählten Kategorie. Sprecher A beispielsweise kategorisiert einen Dialog über Whiskey und Brandy mit der Bezeichnung „Alkohol“, wohingegen Sprecher B die Bezeichnung „Getränke“ oder „Spirituosen“ wählt. Entgegen dem individuellen Charakter der Kategorienwahl beruht die betrachtete Sequenz an sich hingegen auf dem *joint action*-Prinzip. Sie wird demnach gemeinsam von den am Dialog beteiligten Personen aufgebaut und besteht aus mindestens zwei Beiträgen.

4.3.2 Der Themenerkennung

Zusammengefasst definiert die zuvor aufgestellte Definition ein Dialogthema als **Kategorie**, welche die in den Dialogbeiträgen genannten **Konzepte** subsumiert. Der von Chahine et al. (2011) vorgestellte Ansatz, in einem ersten Schritt die *Konzeptualisierung* einzelner Wörter auf Grundlage von Wikipedia-Artikeln und in einem zweiten Schritt die *Kategorisierung* der resultierenden Konzepte auf Grundlage von Wikipedia-Kategorien vorzunehmen, bildet somit eine mit der Definition konforme Realisierung der Themenerkennung. Dieser Ansatz wird daher und aufgrund der in Abschnitt 3.1.3 diskutierten Vorteile der Online-Enzyklopädie **Wikipedia** in das Konzept übernommen. Ausgangspunkt für die technische Umsetzung des Themenerkenners bildet deshalb ein mentales Lexikon, welches zum einen das von der Wikipedia bereitgestellte Konzeptwissen in Form von Artikeln und zum anderen das ebenfalls von der Wikipe-

die bereitgestellte Wissen über thematische Zusammenhänge in Form von Kategorien beinhaltet.

Mit dem Ziel, eine auf diesem grundlegenden Ansatz aufbauende Architektur für den Themenerkennung festzulegen, erfolgt nun die Analyse und Festlegung weiterer Konzeptentscheidungen auf Basis der an den Themenerkennung gestellten Anforderungen und der im vorherigen Kapitel vorgestellten Forschungsarbeiten.

Anforderungsanalyse

Die an ein technisches System gestellten Anforderungen ergeben sich aus dessen Zielanwendung, in dem vorliegenden Fall also aus dessen Einsatz in natürlichen Dialogen zwischen einem menschlichen und einem künstlichen Gesprächspartner. Somit lassen sich folgende Anforderungen an den Themenerkennung zusammenfassen:

- **Generalität** Der Themenerkennung muss generell genug sein, um in nicht-aufgabenorientierten, domänenunabhängigen Dialogen eingesetzt werden zu können.
- **Echtzeitfähigkeit** Der Themenerkennung muss eine Verarbeitung neuer Dialogbeiträge zur Laufzeit der Interaktion gewährleisten, um das System jederzeit mit Informationen über die aktuelle Themensituation ausstatten zu können.

Um aufbauend auf diesen Anforderungen Entscheidungen hinsichtlich der Architektur des Themenerkenners zu treffen, ist eine Abwägung dieser Anforderungen mit dem aktuellen Stand der Forschung unerlässlich. Der nachfolgende Abschnitt verfasst eine entsprechende Diskussion, deren Resultat die Festlegung einer Architektur ist.

Ableitung von Konzeptentscheidungen

Wie in Abschnitt 3.3.2 zusammengetragen, basiert die automatische Erkennung von Themen und die darauf aufbauende Anpassung des Dialogverhaltens bei konversationalen Systemen im Allgemeinen auf entsprechenden Dialogregeln. Die damit verbundenen Nachteile wie der hohe manuelle Aufwand, die Probleme bei der Verarbeitung von nicht in Betracht gezogenen Äußerungen und die auf Ebene der einzelnen Äußerungen beschränkte Themenerkennung veranschaulichten die Notwendigkeit eines von den Regeln losgelösten Erkennungsprozesses. Daher bildet der hier zu entwickelnde Themenerkennung, wie in Abbildung 4.1 dargestellt, eine eigenständige Komponente

innerhalb der Gesamtarchitektur. Die Übertragung der anhand dieser Komponente ermittelten Themeninformationen auf das Zielsystem (in diesem Fall das Dialogsystem des Agenten Max) muss demzufolge über eine Schnittstelle realisiert werden.

Nachdem die Einordnung des Themenerkenners in die Gesamtarchitektur geklärt ist, widmen wir uns nun der Ausarbeitung eines Konzepts, welches den zuvor aufgestellten Anforderungen entspricht. Maas (2007) führte zwei Implementierungskriterien für die Konzeption eines Themenerkenners im Hinblick auf dessen Einsatz in natürlichen Mensch-Maschine-Interaktionen ein. Die Forderung nach *Generalität* verlangt demnach einen **dynamischen** Themenerkennung, welcher unabhängig von der jeweiligen Domäne Informationen über die in dem Dialog angesprochenen Themen ermittelt. Dieses Kriterium wird durch die Umsetzung der Erkennungsprozesse auf Basis der Wikipedia erfüllt, deren enorme Anzahl an Konzeptbeschreibungen und Kategorien nicht auf eine bestimmte Domäne begrenzt ist. Der Forderung nach *Echtzeitfähigkeit* kommt Maas durch einen **online**-fähigen Prozess nach, wodurch eine Erkennung des Themas ohne Zeitverzögerung ermöglicht wird. Dies gelingt durch eine Aktualisierung der Themeninformationen mit jedem neuen Dialogbeitrag.

Abgesehen von den Anforderungen, die der Themenerkennung durch seinen Einsatz in laufenden Dialogen erfüllen muss, liefert die Betrachtungsweise eines Dialogthemas als ein von beiden Dialogpartnern *gemeinsam* durchgeführtes Projekt (s. Definition Dialogthema, Abschnitt 4.3.1) ein weiteres bei der Konzeption zu beachtendes Kriterium. Demnach muss des Weiteren eine **globale** Themenerkennung realisiert werden, um erst dann ein Dialogthema zu identifizieren, wenn es sich über mindestens zwei Dialogbeiträge erstreckt. Zu diesem Zweck wird ein Ähnlichkeitsmaß benötigt, welches *thematische Überschneidungen* zwischen bereits angesprochenen und mit dem neuen Dialogbeitrag hinzukommenden Konzepten ermittelt.

Eine weitere Konzeptentscheidung betrifft die Ermittlung der Themeninformationen, die jederzeit Aufschluss über die aktuelle Themensituation des betrachteten Dialogs geben sollen. Die Themeninformationen müssen demzufolge Angaben über die von den Dialogteilnehmern initiierten oder bereits durchgeführten Aktionen beinhalten, die nach Gardner (1987) (s. Abschnitt 2.2.2) den thematischen Ablauf eines Dialogs prägen. Die im Rahmen des TDT-Projekts (Allan, 2002) definierten, in Abschnitt 3.2.3 vorgestellten Aufgaben zur automatischen Erkennung und Verfolgung von Themen in Radionachrichten, ermöglichen die Identifikation dieser Aktionen. Allerdings

lassen sich die Aufgaben nicht deckungsgleich auf den hier zu konzipierenden Themenkenner übertragen, da diese für eine *offline*-Anwendung spezifiziert wurden. Die vorgenommene Anpassung der TDT-Aufgaben an die im vorliegenden Fall benötigten Themeninformationen wird im nächsten Abschnitt präsentiert. Festzuhalten ist an dieser Stelle, dass bei der technischen Umsetzung des Themenerkenners eine Themenerkennung realisiert werden muss, die *implizit* während der Identifikation des aktuellen Dialogthemas eine Aktualisierung der benötigten Themeninformationen vornimmt.

Ein weiterer, die Anforderung nach *Echtzeitfähigkeit* betreffender Aspekt ist der maschinelle Zugriff auf das in dem künstlichen mentalen Lexikon enthaltene Wikipedia-Wissen. Um eine zeitnahe Antwort bei Anfragen an die Wissensquelle zu erhalten, müssen die von der Wikipedia bereitgestellten Informationen daher indexiert werden (s. Abschnitt 3.1.4).

Architektur

Abbildung 4.2 präsentiert die aus den oben festgelegten Konzeptentscheidungen hervorgegangene Architektur für einen dynamischen, online-fähigen und globalen Themenkenner. Der große Unterschied zu bisherigen Prozessen zur Erkennung des aktuellen Dialogthemas ist die für eine erfolgreiche Themenidentifikation erforderliche Erkennung von **thematischen Überschneidungen** zwischen mindestens zwei Dialogbeiträgen. Auf diese Weise wird eine *globale* Themenerkennung ermöglicht. Dieser Vorgang wird mit jeder neu eintreffenden Äußerung erneut durchgeführt, sodass stets die Aktualität der Informationen über das Dialogthema gewährleistet ist und eine *online-fähige* Themenerkennung erreicht wird. Die Echtzeitfähigkeit des Systems wird darüber hinaus durch eine vorherige Indexierung des für die Durchführung der Erkennungsprozesse benötigten Wikipedia-Wissens und der dadurch erzielten schnellen Antwortzeit sichergestellt. Der auf diese Weise umgesetzte Zugriff auf kollaborativ erstelltes, domänenunabhängiges Konzeptwissen erfüllt zudem das Kriterium der *dynamischen* Themenerkennung.

Des Weiteren wird durch eine Umsetzung der Erkennungsprozesse, welche implizit die Ableitung weiterer Themeninformationen leistet, zum ersten Mal die Grundlage für die Emulation menschlichen Themenbewusstseins in konversationalen Systemen geschaffen. Hierbei liefert der Themenerkenner nicht nur Informationen bezüglich des aktuellen Themas, sondern generiert darüber hinaus Informationen hinsichtlich des

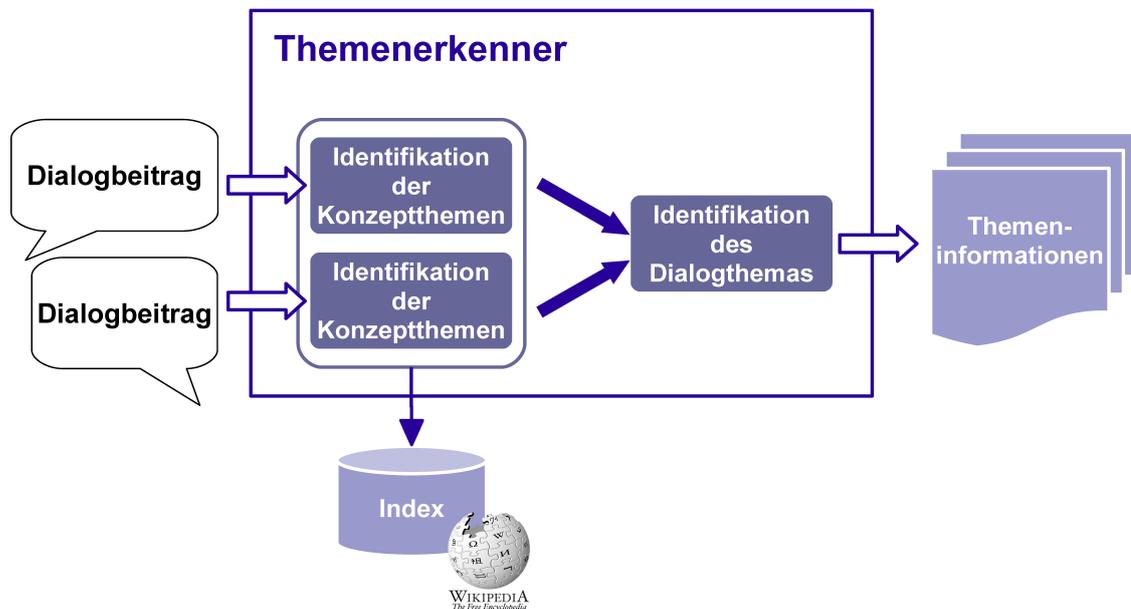


Abbildung 4.2: Architektur des dynamischen, online-fähigen und globalen Themenerkenners, welcher im Wesentlichen aus zwei Komponenten besteht. Auf Grundlage der durch die *Identifikation der Konzeptthemen* ermittelten Informationen über mögliche thematische Richtungen von zwei hintereinander eintreffenden Dialogbeiträgen wird im Rahmen der *Identifikation des Dialogthemas* nach thematischen Überschneidungen gesucht. Darauf aufbauend wird das aktuelle Dialogthema sowie weitere Themeninformationen ermittelt. Das dafür benötigte Wissen wird einem *Wikipedia-Index* entnommen.

thematischen Ablaufs eines betrachteten Dialogs. Auf Basis dieser Informationen gelingt die Generierung entsprechender Reaktionen des Zielsystems auf bestimmte, diesen thematischen Ablauf beeinflussende Aktionen, sowie entsprechender Äußerungen, die diese Aktionen in bestimmten Situationen initiieren beziehungsweise durchführen. Der Ansatz für die genaue Umsetzung dieser Verhaltensweisen innerhalb des Dialogsystems des Agenten Max ist Gegenstand des Abschnitts 4.4. Im Folgenden werden die Themeninformationen sowie deren automatische Ermittlung im Rahmen des Erkennungsprozesses genauer beschrieben.

4.3.3 Themeninformationen

Der thematische Ablauf eines Dialogs, und somit der Aufbau eines gemeinsamen Dialogthemas, wird durch verschiedene, wechselseitige Aktionen der beteiligten Dialogpartner beeinflusst. Ein künstliches Bewusstsein für diese Aktionen erfordert den maschi-

nellen Zugriff auf Informationen über die aktuelle Dialogsituation. Wie im vorherigen Abschnitt festgehalten, soll die Ermittlung dieser zusätzlichen Themeninformationen während der eigentlichen Themenerkennung erfolgen. Für jede dieser Aktionen gilt es daher, eine Aufgabe für den Themenerkennner zu definieren, deren Ausführung in der Spezifikation der entsprechenden Themeninformation resultiert.

Aufgabenspezifikation

Ein Rückblick auf Abschnitt 2.2.2 veranschaulicht, welche Kategorien an thematischen Aktionen vorliegen und inwiefern diese den Themenverlauf eines Dialogs beeinflussen. Nachfolgend erfolgt eine kurze Wiederholung der einzelnen Aktionen sowie die Ableitung entsprechender Aufgaben an den Themenerkennner. Tabelle 4.1 liefert einen Überblick über die thematischen Aktionen, den zugehörigen Aufgaben des Themenerkennners und den daraus resultierenden, stets aktuell zu haltenden Themeninformationen.

Themeneinführung Eine Themeneinführung in natürlichen Dialogen erfolgt genau dann, wenn einer der teilnehmenden Sprecher einen Themenvorschlag macht und sich für beide Dialogteilnehmer daher die Möglichkeit ergibt, ein neues Dialogthema aufzubauen. Nach Schank (1977) ist mit jedem neuen Dialogbeitrag eine Themeneinführung verknüpft, da jede Äußerung Vorschläge für neue Themen enthält. Die Information darüber, ob sich mit dem betrachteten Dialogbeitrag eine Themeneinführung ereignet, ist daher wenig aufschlussreich. Die erste Aufgabe des Themenerkennners besteht demnach in der **Identifikation von Themenvorschlägen**, deren Ergebnis eine Auflistung der eingeführten Themen und somit **möglicher Dialogthemen** ist.

Themenfortführung Wird ein zuvor eingeführtes Thema mit der nächsten Äußerung des Dialogpartners fortgeführt, lässt sich eine thematische Überschneidung zwischen den Dialogbeiträgen und darauf aufbauend ein gemeinsames Dialogthema identifizieren. Die zweite Aufgabe des Themenerkennners ist somit die **Identifikation eines Dialogthemas**. Die Erfüllung dieser Aufgabe resultiert in der Angabe des **aktuellen Dialogthemas**, was gleichzeitig auf die Durchführung einer Themenfortführung hinweist. Lässt sich kein Dialogthema ermitteln, findet zu dem Zeitpunkt auch keine Themenfortführung statt.

Themenwechsel Innerhalb eines laufenden Dialogs ereignet sich ein Themenwechsel genau dann, wenn sich die Dialogpartner kollektiv einem neuen Thema zuwenden und das vorherige aufgeben (Svennevig, 1999). Dieser Wechsel kann auf zwei Arten vollzogen werden: Graduell, in Form eines *Themendriffs*, oder abrupt, in Form eines *Themensprungs*. Die Möglichkeit, mit der nächsten Äußerung einen Themendrift zu vollziehen, ist jederzeit gegeben, da, wie hinsichtlich der Themeneinführung festgehalten, jeder Dialogbeitrag neue Themenvorschläge enthält, die mit der nachfolgenden Äußerung fortgeführt werden können. Die Möglichkeit eines Themensprungs ergibt sich hingegen nur dann, wenn aktuell ein Dialogthema existiert und der aktuelle Dialogbeitrag keine thematische Überschneidung zu diesem Dialogthema aufweist. Die dritte Aufgabe betrifft daher die **Erkennung eines möglichen Themensprungs**. Die tatsächliche Durchführung eines Themenwechsels hängt dann von der darauf folgenden Äußerung des anderen Sprechers und somit von einer anschließenden Themenfortführung ab. Auf eine entsprechende Überprüfung wird an dieser Stelle verzichtet, da sich der letztliche Vollzug eines Themenwechsels (im Gegensatz zu dessen Einleitung) nur selten in der konversationalen Verhaltensweise niederschlägt.

Themenwiedereinführung Weist der aktuelle Dialogbeitrag keine thematische Überschneidung mit der zuvor bearbeiteten, jedoch mit einer anderen vorangegangenen Äußerung auf, spricht man von einer Themenwiedereinführung. Die für eine automatische Wahrnehmung dieser Aktion erforderliche Aufgabe ist daher die **Identifikation thematischer Überschneidungen zu früheren Äußerungen**, welche dann gegebenenfalls mit der Identifikation eines bereits besprochenen Dialogthemas einhergeht.

Themenbenennung Die Themenbenennung hat keinen direkten Einfluss auf den Themenverlauf, liefert jedoch die Möglichkeit, auf ein Thema zu referenzieren und auf diese Weise über eben dieses zu sprechen. Über diese Aktion ist wenig bekannt, Schneider (1988) weist jedoch darauf hin, dass die von Menschen meist bevorzugte nominale Form zur Benennung eines Themas mit dessen Repräsentation innerhalb des *mental*en Lexikons korreliert. Eine entsprechende **Bezeichnung für ein Thema** wird durch Durchführung der **Festlegung einer Themenbezeichnung** gefunden.

Tabelle 4.1: Überblick über die thematischen Aktionen, den zugehörigen Aufgaben des Themenerkenners und den jeweiligen Themeninformationen.

Aktion	Aufgabe	Themeninformation
Themeneinführung	Identifikation von Themenvorschlägen	Mögliche Dialogthemen
Themenfortführung	Identifikation eines Dialogthemas	Aktuelles Dialogthema
Themenwechsel	Erkennung eines möglichen Themensprungs	Ja/Nein
Themenwiedereinführung	Identifikation thematischer Überschneidungen zu früheren Äußerungen	Aktuelles Dialogthema
Themenbenennung	Festlegung einer Themenbezeichnung	Themenbezeichnung

4.4 Integration der Themeninformationen

Wie zu Beginn dieses Kapitels aufgezeigt, muss für eine Emulation menschlichen Themenbewusstseins für den konversationalen Agenten Max neben der Realisierung eines Themenmodells ein zweiter Aufgabenschwerpunkt bearbeitet werden: die Integration der anhand des Themenerkenners ermittelten Themeninformationen in das Dialogsystem des Agenten. Die ursprünglich auf Regeln basierende Realisierung des Dialogsystems gilt es dabei aufgrund ihrer Robustheit zu bewahren. Die Bewältigung dieser Aufgabe unterteilt sich somit in zwei Teilschritte: Erstens, die Einbettung des Themenerkenners in die bestehende Architektur des Dialogsystems, und zweitens, die Anpassung des konversationalen Verhaltens des Agenten hinsichtlich aktueller Themeninformationen.

Der vorliegende Abschnitt beginnt mit der Auflistung der Anforderungen, die mit der Integration der Themeninformationen einhergehen. Im Anschluss daran erfolgt eine Beschreibung des bestehenden Dialogsystems des Agenten Max, um darauf aufbauend einen Ansatz für die Einbettung des Themenmodells in die Architektur dieses Systems zu entwickeln. Danach wird, unter Berücksichtigung bestehender Arbeiten, ein Konzept für die Anpassung des Konversationsverhaltens von Max an die aktuelle Themensituation eines laufenden Dialogs ausgearbeitet.

4.4.1 Anforderungsanalyse

An die Herausforderung, mit der Implementierung eines künstlichen Bewusstseins für Themen in Dialogen zur Verbesserung der intuitiven Mensch-Maschine-Interaktion beizutragen, sind folgende zwei Anforderungen geknüpft:

- **Echtzeitfähigkeit** Die gegenwärtige Echtzeitfähigkeit des Agenten darf durch das zusätzliche Themenbewusstsein nicht beeinträchtigt werden.
- **Natürlichkeit** Das Wissen über Gesprächsthemen muss derartig in das Konversationsverhalten des Agenten einfließen, dass die Ausführung adäquater Äußerungen hinsichtlich der aktuellen Themensituation während der Interaktion mit Menschen gelingt.

Die Forderung nach der Echtzeitfähigkeit betrifft dabei lediglich die Umsetzung des ersten Integrationsschrittes, also der Einbettung des Themenerkenners in die Systemarchitektur, da die bestehenden konversationalen Fähigkeiten diese Anforderung bereits erfüllen. Der Aspekt der Natürlichkeit hingegen bezieht sich auf die Integration der aus der Themenerkennung resultierenden Informationen in das Konversationsverhalten von Max.

4.4.2 Einbettung des Themenerkenners in die Systemarchitektur

Die physikalische Einbettung der aus dem in Abschnitt 4.3.2 konzipierten Themenerkenners hervorgehenden Informationen in das konversationale Verhalten des Agenten Max verlangt ein Verständnis für die Architektur und die Umsetzung des bestehenden Dialogsystems des Agenten. Im Folgenden wird das dafür benötigte Wissen durch eine Beschreibung dieser Komponente vermittelt. Darauf aufbauend gelingt schließlich die Einordnung des Themenmodells in die Architektur des Dialogsystems.

Das Dialogsystem von Max

Das Dialogsystem des Agenten Max wird durch eine **Deliberative Komponente** realisiert. Wie in Abbildung 4.3 veranschaulicht, besteht diese aus vier Modulen, die

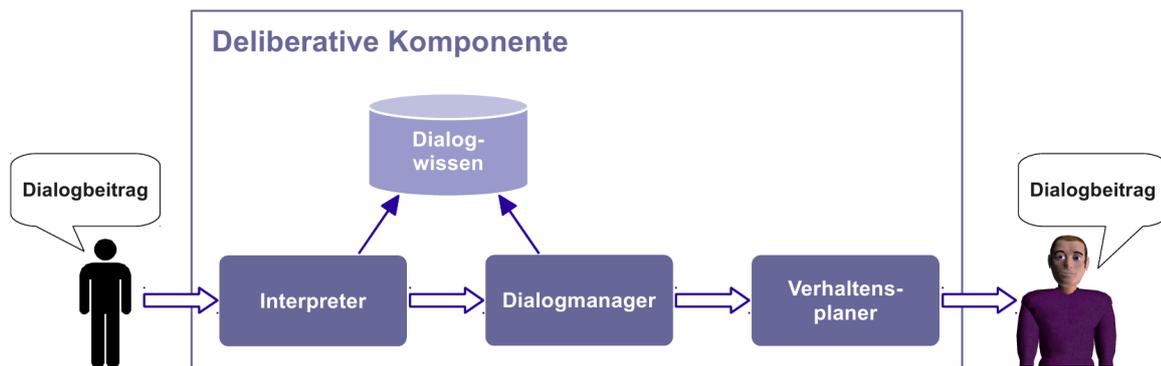


Abbildung 4.3: Die konversationalen Fähigkeiten des virtuellen Agenten Max werden durch die *Deliberative Komponente* ermöglicht.

nacheinander die Eingaben des menschlichen Dialogpartners verarbeiten. In einem ersten Schritt ermittelt der *Interpreter* durch einen Musterabgleich und unter Zugriff auf das *Dialogwissen* die Bedeutung des Eingabetexts. Diese wird in einer Zwischenrepräsentation dargestellt, die verschiedene Ebenen für die Angabe der unterschiedlichen Aspekte der Funktion der Äußerung enthält. Das Ergebnis der Interpreteranalyse wird zusammen mit dem ursprünglichen Text an den *Dialogmanager* übergeben, der, ebenfalls unter Rückgriff auf das Dialogwissen, eine entsprechende Antwort auswählt und diese der *Verhaltensplaner* übergibt. Dort wird die Antwort in eine multimodale Äußerung überführt.

Integration der Themeninformationen in das Dialogwissen

Wie aus Abbildung 4.3 ersichtlich, basieren sowohl die Interpretation der Eingaben des menschlichen Dialogpartners als auch die Auswahl einer entsprechenden Antwort des Agenten auf dem das konversationale Verhalten von Max definierende **Dialogwissen**. Um also die Äußerungen des Agenten an die aktuelle Themensituation des laufenden Dialogs anzupassen, ist eine Einbettung des Themenerkenners erforderlich, die, wie in Abbildung 4.4 dargestellt, eine direkte Integration der Themeninformationen in das Dialogwissen ermöglicht. Dies erfordert die Umsetzung einer Kommunikationsschnittstelle zwischen der Themenerkennungskomponente des Themenmodells und der Deliberativen Komponente des Agentensystems.

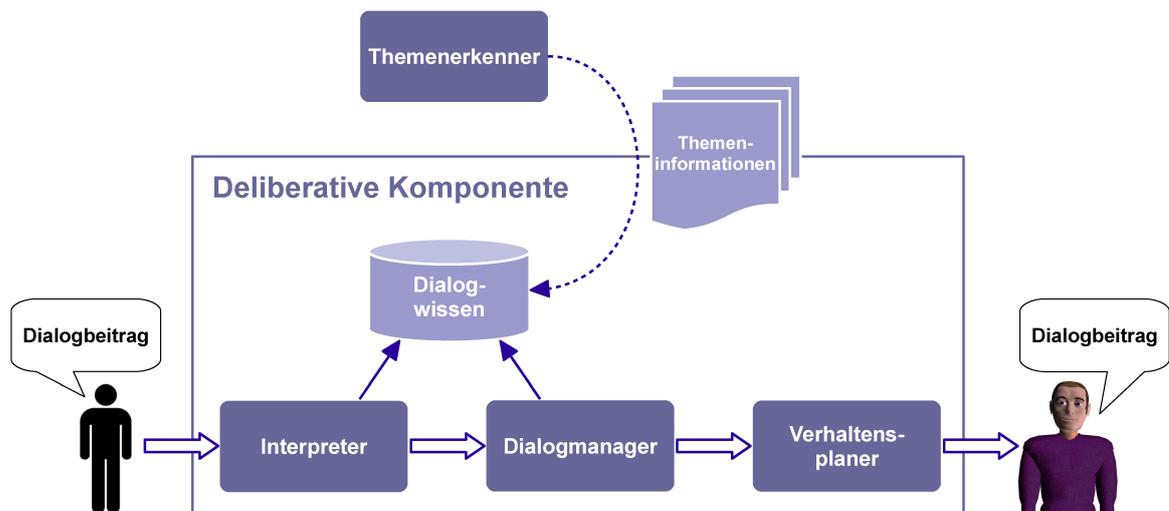


Abbildung 4.4: Integration der Themeninformationen in die Deliberative Komponente.

4.4.3 Themenbasierte Anpassung des konversationalen Verhaltens

Nachfolgend werden unter Berücksichtigung der in den Kapiteln 2 und 3 vermittelten Grundlagen Konzeptentscheidungen diskutiert, die eine die Forderung nach Natürlichkeit erfüllende Integration der Themeninformationen in die Verhaltensweisen von Max realisieren. Darauf aufbauend lassen sich genaue Spezifikationen dieser Verhaltensweisen festhalten.

Ableitung von Konzeptentscheidungen

Die Forderung nach Natürlichkeit verlangt einerseits ein Bewusstsein für die in Abschnitt 2.2.2 aufgeführten Aktionen, die eine **Beeinflussung des thematischen Ablaufs eines Dialogs** zur Folge haben. Anhand dieses Bewusstseins muss Max zum einen *reaktiv* auf solche Aktionen, die seitens des menschlichen Dialogpartners initiiert werden, agieren und zum anderen die Fähigkeit aufweisen, diese Aktionen selbst, das heißt *proaktiv*, durch entsprechende Äußerungen einzuleiten. Auf diese Weise wird das künstliche Themenbewusstsein des Agenten gegenüber seinem menschlichen Dialogpartner demonstriert. Dabei sind sowohl die reaktiven als auch die proaktiven Verhaltensweisen so umzusetzen, dass sie lediglich zu angemessenen Zeitpunkten in Erscheinung treten. Der nachfolgende Abschnitt definiert für jede Aktion eine entsprechende

Verhaltensweise des Agenten.

Natürlichkeit erfordert andererseits auch ein Bewusstsein für die **Angemessenheit eines Themas**, die sich von der jeweiligen Dialogsituation und der persönlichen Beziehung zwischen den Dialogteilnehmern ableiten lassen muss. Die Dialogsituation spiegelt die Umgebung, in welcher der Dialog stattfindet, wider und bezieht dabei sowohl den Grund des Zusammentreffens als auch den kulturellen Hintergrund mit ein. Für die Dialogsituation, die sich aufgrund des gegebenen Anwendungsszenarios des Agenten nicht ändert, lässt sich eine feste Menge an Dialogthemen definieren. Diese Menge an möglichen Dialogthemen muss jedoch bezüglich der persönlichen Beziehung zwischen den Dialogpartnern je nach Bekanntheitsgrad zwischen Max und der jeweiligen Person gegebenenfalls eingeschränkt werden. Genauer gesagt müssen Themen wie Religion oder Politik in Anlehnung an Schneider (1988) (s. Abschnitt 2.2.2) genau dann als „Tabuthemen“ deklariert werden, wenn sich die jeweilige Person zum ersten Mal mit dem Agenten unterhält. Mit wachsender Bekanntschaft zwischen den Dialogpartnern nimmt dann die Auswahl an möglichen Dialogthemen zu.

Definition von Verhaltensweisen

Für die einzelnen, in Abschnitt 4.3.3 aufgeführten Aktionen, die den thematischen Ablauf eines Dialogs beeinflussen (die *Themenbenennung* ist somit ausgeschlossen), sollen an dieser Stelle entsprechende Verhaltensweisen für den Agenten Max festgelegt werden. Diese umschließen sowohl reaktive als auch proaktive Handlungen, sodass der Agent einerseits sein Bewusstsein gegenüber den von seinem menschlichen Gegenüber initiierten Aktionen demonstriert und andererseits dazu in der Lage ist, die einzelnen Aktionen von sich aus, also initiativ, einzuleiten.

Themeneinführung Wie bereits festgehalten, geht jeder neue Dialogbeitrag mit einer Themeneinführung einher. Der Agent führt diese Aktion daher, genau wie sein menschlicher Dialogpartner, mit der Ausführung jeder seiner Äußerungen durch. Eine proaktive Verhaltensweise ist an dieser Stelle somit nicht zu definieren. Die Reaktion des Agenten auf eine von seinem Gegenüber durchgeführte Themeneinführung sollte jedoch je nach vorgeschlagenem Thema unterschiedlich ausfallen. Handelt es sich um die Einführung eines der jeweiligen Dialogsituation angemessenen Themas, muss Max mit einer **Fortführung dieses Themas** reagieren, um kohärente Dialoge zu ermög-

lichen. Als Reaktion auf den Vorschlag eines unangemessenen Themas muss Agent hingegen eine **Empörung** zum Ausdruck bringen, um sein Bewusstsein hinsichtlich der Adäquatheit von Themen zu verdeutlichen.

Themenfortführung Um einen kohärenten Dialog zwischen Max und seinem menschlichen Gesprächspartner zu realisieren und ein Thema über eine längere Dialogsequenz hinweg aufrecht zu erhalten, wird die Fortführung eines Themas fortan als **Ziel des Agenten** definiert. Dementsprechend führt der Agent, wenn angemessen, mit jeder seiner Äußerungen eine Themenfortführung durch. Auf die Fortführung eines Themas seitens seines Dialogpartners zeigt Max keine besondere Reaktion.

Themenwechsel Da jede neue Äußerung neue thematische Richtungen mit sich bringt, ist ein *Themendrift* durch Fortführung eines dieser Richtungen mit jeder Äußerung möglich. Bei einem *Themensprung* hingegen enthält die betrachtete Äußerung keine Überschneidung zu vorherigen Themenvorschlägen. Max muss auf solche Veränderungen mit **Verwunderung** reagieren und auf diese Weise zum Ausdruck bringen, dass ein abrupt eingeführter Themenwechsel für Verwirrung des anderen Dialogteilnehmers sorgen kann. Der Agent selbst soll genau dann einen Themensprung initiieren, wenn das aktuelle **Dialogthema über einen sehr langen Zeitraum aufrecht** erhalten wurde. Dabei ist jedoch darauf zu achten, dass Max den plötzlichen Themenwechsel durch einen entsprechenden Kommentar einführt und sein Gegenüber auf diesen vorbereitet.

Themenwiedereinführung Eine Themenwiedereinführung kann unterschiedliche Reaktionen hervorrufen. Der Agent soll auch hier mit **Verwunderung** reagieren, um dessen Bewusstsein dafür, dass das eingeführte Thema bereits besprochen wurde, zu verdeutlichen. Der Agent selbst sollte genau dann eine Themenwiedereinführung veranlassen, wenn er diesem Thema nachträglich noch **etwas hinzufügen** möchte.

4.5 Zusammenfassung

Auf Basis der in Kapitel 2 und 3 vermittelten Grundlagen und der in der Einleitung formulierten Aufgabenbeschreibung wurde innerhalb dieses Kapitels ein Konzept

für die Realisierung künstlichen Themenbewusstseins in natürlichen Dialogen gearbeitet. Das Gesamtkonzept wurde dabei in zwei Aufgabenschwerpunkte unterteilt, deren Ziele zum einen die Erkennung von Themen in laufenden Dialogen und zum anderen die darauf aufbauende Generierung themenbewussten Konversationsverhaltens des konversationalen Agenten Max sind. Gemäß dieser Unterteilung folgte die Einzeldarstellung dieser Aufgaben und der damit verbundenen Anforderungsanalyse und Festlegung von Konzeptentscheidungen. Als Ergebnis wurden ein Themenmodell für die Umsetzung dynamischer, online-fähiger und globaler Themenerkennung sowie ein Ansatz für die Integration der resultierenden Themeninformationen in das Dialogwissen des Agenten Max konzipiert. Die nachfolgenden Kapitel 5 und 6 beschreiben die technische Umsetzung dieser Konzepte und veranschaulichen die Erweiterbarkeit der konversationalen Fähigkeiten von Max auf Basis künstlichen Themenbewusstseins.

5 Ein Modell zur Erkennung aktueller Dialogthemen

Eine Realisierung künstlichen Themenbewusstseins für konversationale Agenten setzt zum einen die automatische Erkennung von Themen in natürlichsprachlichen Dialogen und zum anderen die Integration der resultierenden Themeninformationen in das bestehende Agentensystem voraus. In Abschnitt 4.3 wurde ein Themenmodell zur Bewältigung der ersten dieser beiden Aufgabenschwerpunkte konzipiert. Abbildung 5.1 liefert einen Überblick über die Gesamtarchitektur dieses Modells. Demnach erhält der **Themenerkennner** sukzessive die in Textform vorliegenden Dialogbeiträge, verarbeitet diese unter Zugriff auf den **Wikipedia-Index** und leitet die resultierenden **Themeninformationen** an ein Zielsystem, hier das Dialogsystem des konversationalen Agenten Max, weiter.

Das vorliegende Kapitel beschäftigt sich ausführlich mit der technischen Umsetzung des Themenmodells. Daraus ergibt sich folgender Aufbau dieses Kapitels: Abschnitt 5.1 widmet sich dem Zugriff auf Informationen über Konzepte und ihre Zusammenhänge und beschreibt die dafür erforderliche Definition maschinenlesbaren Konzeptwissens auf Grundlage der Online-Enzyklopädie Wikipedia. Die daran anschließenden Abschnitte beschäftigen sich mit der technischen Umsetzung des Themenerkennners. Dabei wird in Abschnitt 5.2 zunächst ein kurzer Überblick über die im vorherigen Kapitel zusammengetragenen Konzeptentscheidungen geliefert, welche sich aus dem Einsatz des Themenerkennners in natürlichen Dialogen zwischen einem Menschen und einem konversationalen Agenten ergeben. Anschließend werden in Abschnitt 5.3 die einzelnen Komponenten des Themenerkennners vorgestellt, die eine an die zuvor aufgestellten Kriterien angepasste Umsetzung der Aufgaben zur Spezifikation der Themeninformationen ermöglichen. Abschnitt 5.4 befasst sich daraufhin mit der Evaluation des automatischen Themenerkennners auf Basis thematisch vorsortierter Zeitungsartikel und dem von Endrass, Rehm und André (2011) erstellten *CUBE-G*-Korpus. Das Kapitel schließt mit einer Zusammenfassung der wichtigsten Ergebnisse.

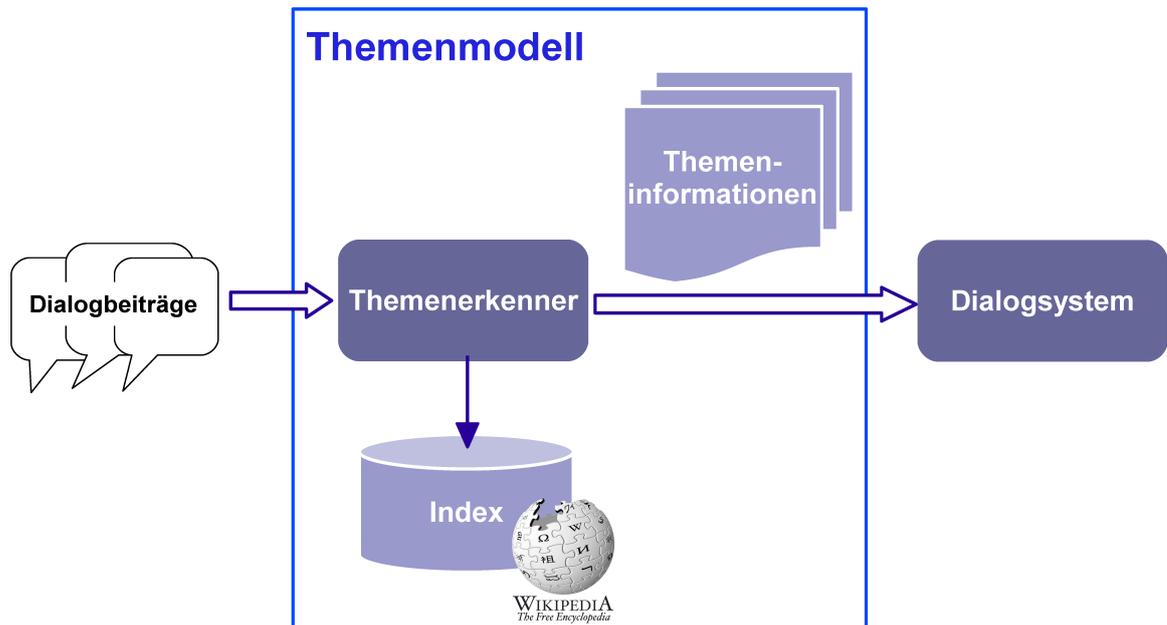


Abbildung 5.1: Gesamtarchitektur des Themenmodells. Auf Grundlage der im betrachteten Dialog geäußerten *Beiträge*, die sukzessive an den *Themenerkennung* übermittelt werden, und unter Zugriff auf den Wikipedia-basierten *Index* werden *Informationen hinsichtlich der aktuellen Themensituation* ermittelt und an das nachfolgende System weitergeleitet.

5.1 Wikipedia als Quelle konzeptueller Wissensstrukturen

Laut der in Abschnitt 4.3.1 ausgearbeiteten Definition ist ein Dialogthema eine **Kategorie**, welche die in einer Sequenz an Dialogbeiträgen genannten Konzepte subsumiert und generalisiert. Der Kategoriengraph von Wikipedia stellt zahlreiche Kategorien bereit, welche die in den natürlichsprachlichen Artikeln beschriebenen Konzepte subsumieren. Der Themenerkennung macht Gebrauch von dieser Analogie und identifiziert das Dialogthema durch die Abbildung der einzelnen, in den Dialogbeiträgen enthaltenen Terme auf Wikipedia-Artikel und durch die daran anschließende Spezifikation ihrer gemeinsamen Wikipedia-Kategorien. Auf diese Weise bezieht der Themenerkennung nicht nur die in den Äußerungen genannten Wörter in den Erkennungsprozess mit ein, sondern spezifiziert ein Dialogthema auf Grundlage umfangreichen Konzeptwissens, welches kollaborativ von zahlreichen freiwilligen Autoren erstellt, gewartet und erweitert wird. Das in der Online-Enzyklopädie Wikipedia enthaltene Wissen wird in-

nerhalb des Themenmodells somit als künstliches *mentales Lexikon* (s. Abschnitt 2.3) genutzt. Im Folgenden werden die dafür benötigten, von der Wikipedia bereitgestellten Informationen genauer spezifiziert und die maschinenlesbare Darstellung dieser Informationen vorgestellt.

5.1.1 Informationsspezifikation

Wie bereits in Abschnitt 3.1.2 erwähnt, lässt sich das in Wikipedia enthaltene Wissen als *teilstrukturiert* bezeichnen. Genauer gesagt werden die Inhalte der Wikipedia als ein Hypertext organisiert, dessen Syntax einfache Anweisungen für die Textformatierung und das Anlegen von Querverweisen in Form von Hyperlinks erlaubt. Anhand dieser Syntax lassen sich Strukturelemente ermitteln, die Informationen hinsichtlich der in den einzelnen Artikeln beschriebenen Konzepte und ihrer Zusammenhänge ableiten lassen. Das Themenmodell zieht einige dieser Informationen in den Erkennungsprozess mit ein, die im Folgenden genauer vorgestellt werden.

Artikeltexte Artikeltexte dienen der natürlichsprachlichen Beschreibung einzelner Konzepte. Das Themenmodell nutzt die Artikeltexte daher als Menge an Begriffen, die mit dem jeweiligen Konzept in Verbindung stehen, und führt auf diesen eine textbasierte Suche nach einem passenden Artikel für einen gegebenen Term aus.

Artikeltitel Jeder einzelnen Seite in Wikipedia wird ein eindeutiger Titel zugewiesen, welcher zur Identifikation der jeweiligen Seite dient. Demnach besitzt auch jeder Artikel in Wikipedia einen solchen Titel. Das Themenmodell nutzt diese Artikeltitel als Bezeichnungen für die in den Artikeln beschriebenen Konzepte.

Eingehende Querverweise Mittels interner Hyperlinks können Artikel auf andere Artikel verweisen, um so eine Relation zwischen den darin beschriebenen Konzepten herzustellen und die Benutzer bei der Navigation durch die zahlreichen Artikel zu unterstützen. Der in Abbildung 5.2 dargestellte Ausschnitt des Artikels über den virtuellen Agenten Max enthält beispielsweise mehrere Querverweise auf andere Artikel, welche an der blauen Hervorhebung der entsprechenden Textstelle innerhalb des Artikeltexts zu erkennen sind (existiert der Zielartikel in der Wikipedia noch nicht, wird die entsprechende Textstelle rot hervorgehoben). Bei dem Querverweis auf den Artikel über die Universität Bielefeld stimmt der Titel

des verlinkten Artikels mit dem Textschnipsel, der auf diesen Artikel verweist, überein. Oftmals bedarf es jedoch einer Anpassung des ursprünglichen Artikeltitels an den den jeweiligen Querverweis umgebenden Text, wenn zum Beispiel die Pluralform oder der Genitiv eines Konzepts benötigt oder aber die Lesbarkeit durch eine andere Konzeptbezeichnung erhöht wird. So verweist der abgebildete Artikel über den Agenten Max zum Beispiel auf den Artikel „*Software-Agent*“, der diesem Querverweis hinterlegte Text lautet jedoch „*Agent*“ (die entsprechende Textstelle ist in der Abbildung ebenfalls rot umrahmt). Der Themenerkennner nutzt die Textschnipsel eingehender Querverweise als Synonyminformation für den jeweils dahinter verlinkten Artikel.

Avatar Max

Max (the **M**ultimodal **A**ssembly **eX**pert) ist ein virtuelle **Agent**, der unter anderem als Museumsführer^[1] im **Heinz Nixdorf MuseumsForum** in **Paderborn** eingesetzt wird^[2]. „**Avatar**“ wird in diesem Zusammenhang verstanden als sichtbare Erscheinung einer unsichtbaren, eigenständig agierenden Software-Maschine, verkörpert als eine computeranimierte Figur. Auf Grundlage einer **BDI-Architektur** verfügt Max über Fähigkeiten, die es ihm erlauben, Menschen anzusprechen und in ein Gespräch zu verwickeln. Dafür greift der Agent derzeit auf über 1.200 bestehende Regeln seiner Wissensbasis zu. Mit Hilfe einer Kamera nimmt Max seine Umgebung wahr und reagiert auf diese. Er kann zudem ihm bekannte Gesichter erkennen und sein Gegenüber entsprechend ansprechen.

Max wurde ab 1999 im Rahmen der Dissertation von **Stefan Kopp** in der von **Ipke Wachsmuth** geleiteten Arbeitsgruppe Wissensbasierte Systeme/**Künstliche Intelligenz**^[3] der **Universität Bielefeld** erschaffen und dort in vielen Arbeiten weiterentwickelt. Heute wird er insbesondere als Forschungsplattform im **Sonderforschungsbereich** "Alignment in Communication" (SFB 673) und im Center of Excellence "Cognitive Interaction Technology" (CITEC) eingesetzt.

Abbildung 5.2: Ausschnitt des Wikipedia-Artikels über den virtuellen Agenten Max. Hervorgehoben sind zwei Querverweise auf andere Wikipedia-Artikel, deren hinterlegte Textschnipsel als Synonyminformation verwendet werden.

Weiterleitungsseiten Eine weitere Quelle für Synonyminformationen liefern die sogenannten Weiterleitungsseiten der Wikipedia. Diese Seiten haben keinen Inhalt, sondern leiten den Benutzer anhand des von ihm eingegebenen Suchbegriffs zu

einem Artikel weiter, der das dem jeweiligen Begriff zugrunde liegende Konzept beschreibt. Die Titel der Weiterleitungsseiten zeigen demnach weitere gängige Bezeichnungen für Konzepte auf und werden ebenfalls als Synonyme für die jeweiligen Konzepte interpretiert.

Kategorien Die ebenfalls kollaborativ definierten Kategorien der Wikipedia ermöglichen eine thematische Klassifikation der Artikel und somit der darin beschriebenen Konzepte. Der Themenerkennung verwendet die Kategorien daher als Menge möglicher Dialogthemen.

Die implizit von der Wikipedia bereitgestellten Informationen beruhen auf der kollaborativen Leistung zahlreicher freiwilliger Autoren und liefern somit eine Repräsentation der allgemeinen menschlichen Auffassung von Konzepten und ihren Zusammenhängen (s. Abschnitt 3.1.3). Diese Betrachtungsweise bildet das Fundament für den zuvor erwähnten Aufbau eines künstlichen, maschinenlesbaren mentalen Lexikons auf der Grundlage der Online-Enzyklopädie, welches das für eine automatische Erkennung von Dialogthemen und den den thematischen Ablauf eines Dialogs beeinflussenden Aktionen benötigte Konzeptwissen bereitstellt. Die genaue Realisierung eines maschinenlesbaren Zugriffs auf das in Wikipedia enthaltene Konzeptwissen ist Gegenstand des folgenden Abschnitts.

5.1.2 Informationsindexierung

In Abschnitt 3.1.4 wurde die **Indexierung** als Ansatz für einen maschinellen Zugriff auf die von Wikipedia bereitgestellten Wissensstrukturen vorgestellt, welcher aufgrund seiner raschen Antwortzeit die geforderte Echtzeitfähigkeit des Erkennungsmoduls unterstützt. Voraussetzung dafür ist das Vorhandensein einer Repräsentation der benötigten Wissensstrukturen in einem maschinenlesbaren Format. Wie in Abschnitt 3.1.3 motiviert, bietet diesbezüglich die direkte Nutzung der von der *Wikimedia Foundation* zum freien Download freigegebenen Wikipedia-Inhalte gegenüber einer Anbindung an die *DBpedia* den elementaren Vorteil der Aktualität. Den Benutzern stehen dabei unterschiedliche Dateien zum Download zur Verfügung, die verschiedene Teile der Wikipedia in XML-Format bereitstellen. Zu diesem Zweck wurde die neueste Version der Datei mit allen deutschen Wikipedia-Artikeln von der Wikimedia-Website¹

¹<http://dumps.wikimedia.org/>

heruntergeladen. Um dieser anschließend die für den Themenerkennung entscheidenden Informationen (s. vorheriger Abschnitt) zu entnehmen, werden alle in der Datei enthaltenen Artikelbeschreibungen nach bestimmten, die gewünschten Informationen kennzeichnenden Strukturelementen durchsucht. Querverweise werden in der Wikipedia-Syntax durch doppelte eckige Klammern definiert (z. B. [[Universität Bielefeld]] oder [[Software-Agent|Agent]], wobei „Agent“ in diesem Fall den Textschnipsel des Hyperlinks darstellt und „Software-Agent“ die genaue Bezeichnung des Artikels, auf welchen verlinkt wird, angibt.

Der Zugriff auf das in dem mentalen Lexikon des Themenmodells enthaltene Konzeptwissen der Wikipedia erfolgt im Zusammenhang mit zwei unterschiedlichen Erkennungsprozessen: Zum einen bei der automatischen Abbildung der in den Dialogbeiträgen angesprochenen Konzepte auf Wikipedia-Artikel und zum anderen bei der Ermittlung der jeweiligen Kategorienzugehörigkeiten für die zuvor identifizierten Artikel (und somit der Themenzugehörigkeiten der genannten Konzepte). Dementsprechend wurden zwei verschiedene Indexe angelegt, die jeweils die für die beiden, im weiteren Verlauf dieses Kapitels (s. Abschnitt 5.3) ausführlicher beschriebenen Prozesse relevanten Wikipedia-Strukturen beinhalten. Beide Indexe wurden auf Basis von *Apache Lucene* (McCandless et al., 2010) erstellt, eine frei verfügbare, auf der Programmiersprache *Java* aufsetzende Programmbibliothek zur Volltextsuche. Der erste Index, **Artikelindex** genannt, vereint die eine Stichwort-basierte Suche nach einem Artikel unterstützenden Informationen. Somit definiert er für jeden innerhalb der Wikipedia angelegten Artikel ein Dokument, welches die zugehörige Wikipedia-interne ID (Identifikationsnummer), den jeweils zugehörigen Titel, die textuelle Beschreibung in Form des Artikeltextes, die eingehenden Querverweise und die Weiterleitungsseiten enthält. Der zweite Index, **Kategorienindex** genannt, beschreibt die von den zahlreichen Autoren der Wikipedia vorgenommene Zuordnung der einzelnen Artikel zu Kategorien. Dementsprechend definiert dieser Index für jede Wikipedia-Kategorie ein Dokument, welches den Titel der betrachteten Kategorie und die Titel der mit dieser Kategorie direkt oder indirekt, das heißt über etwaige Unterkategorien, verknüpften Artikel beinhaltet. Artikel, die mehrfach mit einer Kategorie zusammenhängen, sind entsprechend oft in dem jeweiligen Dokument des Index enthalten, um dessen Relevanz innerhalb des mit der Kategorie beschriebenen Themengebiets zu erhöhen. Die Tabellen 5.1 und 5.2 veranschaulichen den inhaltlichen Aufbau der beiden Indexe.

Tabelle 5.1: Auszug des Index, der zur Abbildung von Termen auf Wikipedia-Artikel genutzt wird.

ID	Titel	Text	Eingehende Querverweise	Weiterleitungsseiten
491828	Universität Bielefeld	Die Universität Bielefeld ist eine 1969 gegründete deutsche Campus-Universität ...	Universitäten Bielefeld, Pädagogischen Hochschule Bielefeld, ...	FH Bielefeld
113274	Ipke Wachsmuth	Ipke Wachsmuth (*27. März 1950) ist ein deutscher Forscher auf den Gebieten ...	Wachsmuth, Ipke	-
...

Tabelle 5.2: Auszug des Index, welcher alle direkten und indirekten Artikel-Kategorie-Zugehörigkeiten enthält.

Kategorie (Titel)	Artikel (Titel)
Politik	Politik, Angela Merkel, Demokratie, Politische Partei, Angela Merkel, Politik, Politik, ...
Sport	Fußball, Dirk Nowitzki, Schiedsrichter, Sport, Fußball, Basketball, ...
...	...

Das durch die beiden Indexe repräsentierte mentale Lexikon des Themenmodells lässt sich von nun an durch den Download und die anschließende Indexierung der aktuellsten Wikipedia-Version auf den neuesten Stand bringen. Der folgende Abschnitt beschreibt, inwiefern und zu welchem Zweck auf das Modell-interne mentale Lexikon zugegriffen wird.

5.2 Dynamische, online-fähige und globale Themenerkennung

Nach der Analyse der Anforderungen und der Spezifikation entsprechender Entwurfskriterien in Abschnitt 4.3.2, fasst dieser Abschnitt das als Resultat aus den genannten

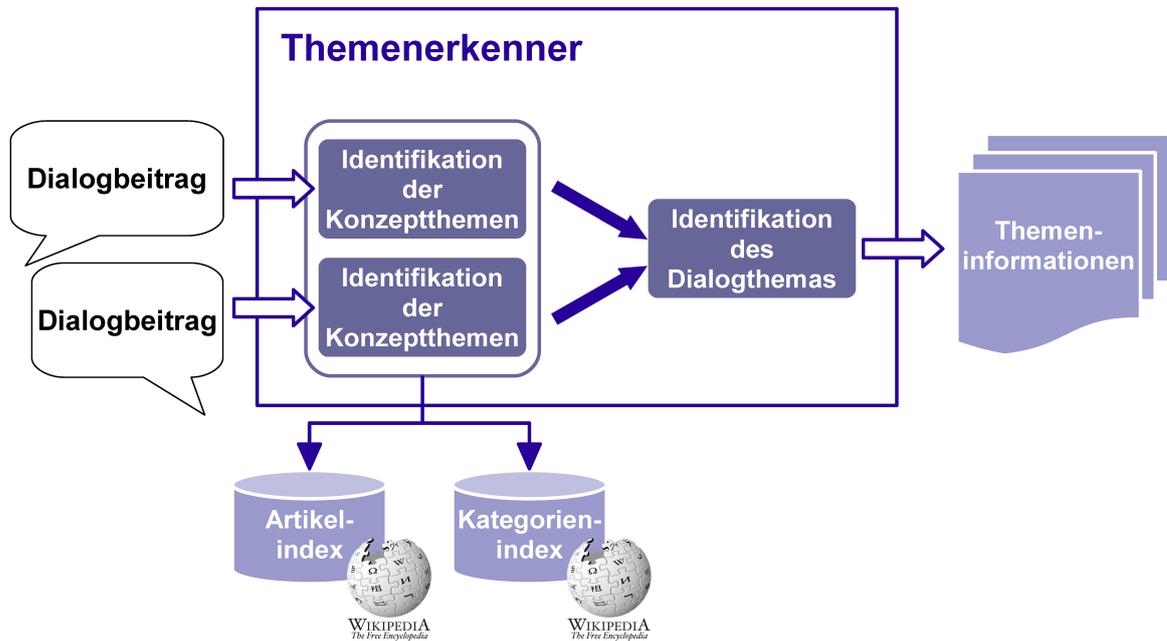


Abbildung 5.3: Basisprozesse zur Identifikation von Dialogthemen.

Vorüberlegungen hervorgegangene Konzept für eine automatische Erkennung von Themen in natürlichen Dialogen zusammen. Die Erkennung von Dialogthemen wird durch die **Themenerkennung**-Komponente des Themenmodells realisiert. Abbildung 5.3 zeigt die im Rahmen der Konzeption entwickelte Architektur des Themenerkenners, auf dessen Grundlage eine *dynamische, online-fähige* und *globale* Themenerkennung gelingt. Die Erkennung eines Dialogthemas auf Basis von mindestens zwei Dialogbeiträgen unterteilt sich dabei im Wesentlichen auf folgende zwei Basisprozesse:

1. Identifikation der Konzeptthemen
2. Identifikation des Dialogthemas

Der erste Prozess des Themenerkenners, die **Identifikation der Konzeptthemen**, spezifiziert unter Zugriff auf das indexierte Wikipedia-Wissen mögliche thematische Richtungen, die mit dem aktuell betrachteten Dialogbeitrag vorgeschlagen werden. Erst durch Fortführung einer dieser Themenvorschläge im weiteren Verlauf des Dialogs lässt sich ein Dialogthema bestimmen. Für den zweiten Prozess, die **Identifikation des Dialogthemas**, bedarf es daher erst einmal einer Überprüfung, ob sich zwei aufeinander folgende Äußerungen hinsichtlich eines Themas überschneiden. Die

Identifikation der Konzeptthemen muss also mindestens zweimal, das heißt auf Basis von zwei separaten Äußerungen, durchgeführt worden sein, bevor eine *thematische Überschneidung* zwischen diesen Äußerungen festgestellt werden kann. Ist diese Bedingung erfüllt, lässt sich das aktuelle Dialogthema durch Ermittlung desjenigen Themas bestimmen, hinsichtlich dessen sich die beiden Äußerungen überschneiden.

Mit jeder neu eintreffenden Äußerung wird der beschriebene Prozessablauf wiederholt. Auf diese Weise gelingt eine **online**-fähige Themenerkennung. Die Identifikation des Dialogthemas auf Basis von zwei aufeinander folgenden Äußerungen entspricht der Betrachtung eines Themas als gemeinsames Projekt beider Dialogteilnehmer und realisiert die geforderte **globale** Erkennung von Dialogthemen. Durch die Umsetzung der beiden Prozesse auf Grundlage des durch die deutsche Wikipedia bereitgestellten Konzeptwissens ist die Themenerkennung darüber hinaus nicht auf eine bestimmte Domäne festgelegt, was schließlich den **dynamischen** Aspekt verwirklicht.

Mit dem Ziel, durch Ausführung der beiden Prozesse genau die für eine Emulation menschlichen Themenbewusstseins erforderlichen Themeninformationen zu erschließen, wurden in Abschnitt 4.3.3 folgende fünf Aufgaben definiert, die der Themenerkennner während der Verarbeitung neuer Dialogbeiträge bewältigen muss:

1. **Identifikation von Themenvorschlägen**
2. **Identifikation eines Dialogthemas**
3. **Erkennung eines möglichen Themensprungs**
4. **Identifikation thematischer Überschneidungen zu früheren Äußerungen**
5. **Festlegung einer Themenbezeichnung**

Der Themenerkennner führt diese Aufgaben implizit während des Prozessablaufs zur Verarbeitung von Dialogbeiträgen aus. Inwiefern die thematischen Informationen anschließend an das Dialogsystem des Agenten Max weitergeleitet und in dessen konversationalen Fähigkeiten integriert werden, ist Gegenstand des nächsten Kapitels. Im Folgenden wird eine genaue Beschreibung der Umsetzung der beiden Basisprozesse geliefert. Danach wird auf die mit der jeweiligen Prozessdurchführung ermöglichten Aufgabenausführungen eingegangen und das Format der resultierenden Themeninformationen vorgestellt.

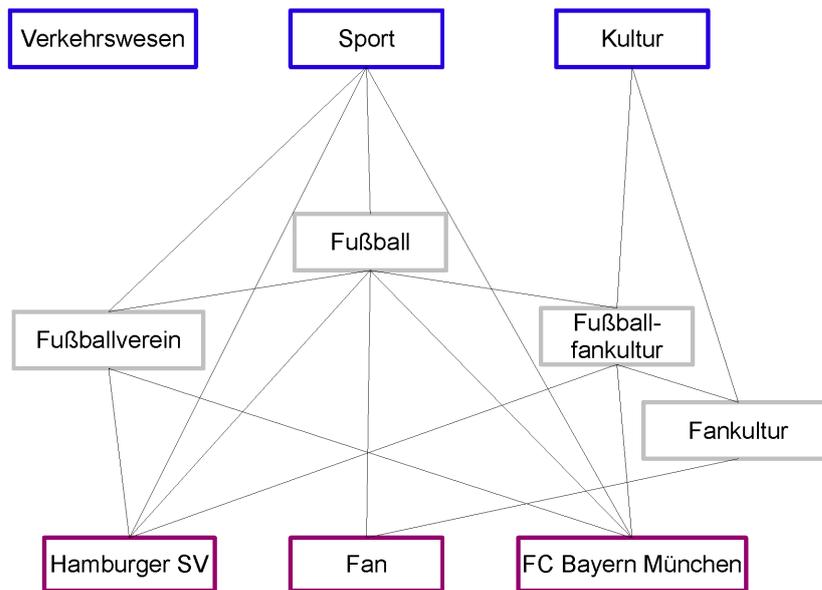


Abbildung 5.4: Konzeptzusammenhänge innerhalb des Wikipedia-Kategoriengraphen.

5.3 Wikipedia-basierte Umsetzung des Themenerkenners

Wie bereits mehrfach dargelegt, wird die technische Umsetzung des zuvor eingeführten Konzepts für eine dynamische, online-fähige und globale Themenerkennung auf Basis des von der deutschen Wikipedia bereitgestellten Konzeptwissens durchgeführt. Verdeutlichen lässt sich der dabei zugrunde liegende Ansatz anhand eines kleinen Beispiels. Gegeben sei der in Abbildung 5.4 dargestellte Ausschnitt des Wikipedia-Kategoriengraphen und ein kurzer Dialog, der aus folgenden zwei Äußerungen besteht:

A: „*Bist du ein Fan von Bayern München?*“

B: „*Nein, eigentlich nicht. Ich mag den HSV.*“

Die einzelnen, in den beiden Äußerungen angesprochenen Konzepte lassen sich auf Wikipedia-Artikel (innerhalb der Abbildung repräsentiert durch die roten Kästchen) „Fan“, „FC Bayern München“ und „Hamburger SV“ abbilden. Der Zusammenhang zwischen diesen Konzepten wird anhand ihrer direkten und indirekten Kategorien-zugehörigkeiten (Kanten zwischen den Kästchen) verdeutlicht. Die Kategorie „Sport“

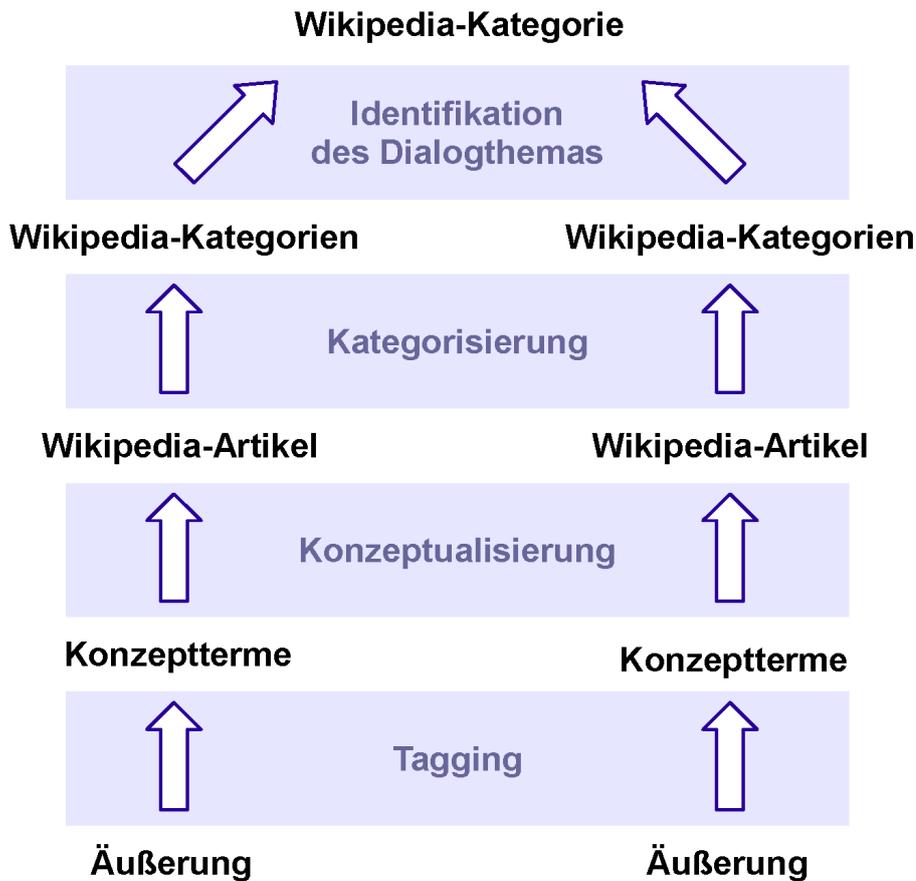


Abbildung 5.5: Überblick über den Gesamtprozess der automatischen Themenerkennung. In der Mitte sind die einzelnen Prozessschritte aufgeführt, deren Ergebnisse die Grundlage für den jeweils nachfolgenden Prozessschritt liefern.

verzeichnet dabei die meisten Zugehörigkeiten und lässt sich somit als gemeinsame Kategorie der genannten Konzepte identifizieren. Der Themenerkennner interpretiert eine bestimmte Menge an Wikipedia-Kategorien (repräsentiert durch die blauen Kästchen) als mögliche Themen eines Dialogs. Die automatische Identifikation eines Dialogthemas geschieht demnach durch die Identifikation derjenigen Wikipedia-Kategorie, welche die in der betrachteten Dialogsequenz genannten Konzepte in Form von Wikipedia-Artikeln subsumiert und zudem einer vordefinierten Menge an Themen-Kategorien angehört. Abbildung 5.5 liefert einen Überblick über die daraus resultierenden einzelnen Prozessschritte des Themenerkenners. Wie im Folgenden vorgestellt, werden diese mit jedem neu eintreffenden Dialogbeitrag durchgeführt, um die Aktualität des dabei ermittelten Dialogthemas zu jedem Zeitpunkt zu gewährleisten.

5.3.1 Identifikation der Konzeptthemen

Ziel des ersten Prozesses ist die Spezifikation möglicher thematischer Richtungen, die mit dem aktuell betrachteten Dialogbeitrag vorgeschlagen werden. Diese bilden die Grundlage für die anschließende Ermittlung thematischer Überschneidungen zwischen den einzelnen Äußerungen und darauf aufbauend die Grundlage für die Identifikation eines Dialogthemas. Die Erkennung von Themenvorschlägen gelingt hierbei durch eine Abbildung der in der jeweiligen Äußerung enthaltenen **Konzeptterme** auf entsprechende Wikipedia-Artikel und der anschließenden Identifikation der Kategorien dieser Artikel. Dies entspricht dem von Chahine et al. (2011) vorgestellten Ansatz, in einem ersten Schritt die **Konzeptualisierung** einzelner Wörter auf Grundlage von Wikipedia-Artikeln und in einem zweiten Schritt die **Kategorisierung** der resultierenden Konzepte auf Grundlage von Wikipedia-Kategorien vorzunehmen.

Konzeptualisierung

Innerhalb der vorliegenden Arbeit besteht ein Konzeptterm aus mindestens einem Nomen oder einem Eigennamen. Die Identifizierung dieser Terme wird durch eine Vorverarbeitung des aktuellen Dialogbeitrags anhand des *Stanford Part-Of-Speech-Tagger* (Toutanova & Manning, 2000) durchgeführt, welcher alle in dem aktuellen Dialogbeitrag enthaltenen Nomen und Eigennamen identifiziert (s. Beschreibung zu Part-of-Speech Tagging in Abschnitt 3.1.4). Des Weiteren extrahiert das System die von dem Tagger erkannten Verben, da auch diese in ihrer jeweiligen Substantivform auf Wikipedia-Artikel abgebildet werden können und auf diese Weise Konzeptinformationen sowie mögliche Themenvorschläge bereitstellen. Dabei werden jedoch ausschließlich Verben betrachtet, die nicht in einer *Stoppwortliste* (s. Abschnitt 3.1.4) enthalten sind, um zum Beispiel Hilfsverben wie „haben“ oder „sollen“ von der Konzeptualisierung auszuschließen. Unter der Annahme, dass die Substantivformen der meisten Verben mit der Infinitivform derselben übereinstimmen, wie beispielsweise für das Verb „schwimmen“ gegeben, erfolgt die Substantivierung mit Hilfe einer weiteren Datenbank. Diese beinhaltet alle in dem Wikiwörterbuch *Wiktionary* enthaltenen konjugierten Verben und ihre zugehörigen Infinitivformen. Existiert für die auf Grundlage dieser Datenbank ermittelte Infinitivform ein gleichnamiger Wikipedia-Artikel, wird diese Infinitivform als Konzeptterm für das betrachtete Verb übernommen. Kann kein gleichnamiger Artikel gefunden werden, wird das jeweilige Verb nicht weiter mit in

den Erkennungsprozess einbezogen.

Die Abbildung der nun identifizierten Konzeptterme des Dialogbeitrags auf die Menge aller Wikipedia-Artikel A_{wiki} erfolgt schließlich durch eine Abbildungsfunktion $f_1 : cterm \mapsto A_{wiki}$, wobei $cterm$ einen der zuvor ermittelten Konzeptterme repräsentiert. Realisiert wird f_1 auf Basis des in Abschnitt 5.1.2 beschriebenen Artikelindex. $cterm$ wird also auf die Menge aller in dem Artikelindex enthaltenen Dokumente D_{art} abgebildet, wobei jedes einzelne Dokument $dArt \in D_{art}$ einen Artikel $art \in A_{wiki}$ repräsentiert:

$$f_1 : cterm \mapsto D_{art} \text{ mit } \forall art \in A_{wiki} \exists dArt \in D_{art}$$

Um anhand dieser Abbildungsfunktion eine entsprechend ihrer Ähnlichkeit zum Konzeptterm $cterm$ geordnete Liste an Wikipedia-Artikeln zu erhalten, wird jedes in dem Artikelindex enthaltene Dokument $dArt \in D_{art}$ auf Grundlage der Ähnlichkeitsfunktion $score(cterm, dArt)$ mit $cterm$ abgeglichen. Somit geschieht die Konzeptualisierung anhand folgender Funktion:

$$f_1(cterm) = \{dArt_1, dArt_2, \dots, dArt_n\} \text{ mit } score(cterm, dArt_1) \geq score(cterm, dArt_2) \geq \dots \geq score(cterm, dArt_n) \quad (5.1)$$

$score(cterm, dArt)$ entspricht dabei der folgenden *Lucene*-internen Formel zur Bewertung der Ähnlichkeit zwischen dem in der Abfrage q spezifizierten Konzeptterm und dem jeweiligen Dokument d :

$$score(q, d) = \sum_{t \in q} (tf(t \in d) \cdot idf(t) \cdot boost_f \cdot norm(q, d)) \quad (5.2)$$

wobei $tf(t \in d)$ die Termfrequenz jedes einzelnen Terms $t \in cterm$ in d spezifiziert und $idf(t)$ die allgemeine Bedeutung von t bezogen auf alle Dokumente aufzeigt. Des Weiteren führt der Verstärkungsfaktor $boost_f$ zu einer Erhöhung des Ähnlichkeitswertes für das betrachtete Dokument bei einer exakten Übereinstimmung von $cterm$ mit dem Titel des jeweiligen Artikels. $norm(q, d)$ kombiniert *Lucene*-interne Normalisierungsfaktoren².

²Für eine genaue Beschreibung der einzelnen Komponenten sei der Leser auf McCandless und Kollegen (2010) verwiesen.

Um auch Konzeptterme zu erkennen, die aus mehreren Wörtern bestehen, wird f_1 im Falle mehrerer hintereinander auftretender Nomen oder für Adjektiv-Nomen-Paare nicht nur mit jedem einzelnen dieser Wörter, sondern zusätzlich mit der Kombination der zusammenhängenden Terme ausgeführt. Das Ergebnis mit dem höheren Ähnlichkeitswert bestimmt die finale Formation des Konzeptterms, das heißt, ob die enthaltenen Terme zusammen einen Konzeptterm formen, oder ob für jeden Term ein einzelner Konzeptterm festgelegt wird. Auf diese Weise fungiert Wikipedia als *Konzeptidentifikator*.

Als Resultat der Konzeptualisierung einer Äußerung wird also für jeden in der Äußerung enthaltenen Konzeptterm eine nach Ähnlichkeit geordnete Liste an Wikipedia-Artikeln zurückgeliefert, welche durch Dokumente des Artikelindex repräsentiert werden. Für den weiteren Verlauf des Erkennungsprozesses werden dabei die obersten zehn Artikel der resultierenden Artikelliste, also die zehn Artikel, die dem betrachteten Konzeptterm am ähnlichsten sind, übernommen und als Menge an Konzepten C_{utt} des betrachteten Dialogbeitrags definiert.

Beispiel Die einzelnen Schritte der Konzeptualisierung lassen sich anhand des in Abbildung 5.6 dargestellten Beispiels zusammenfassen. Für die Äußerung „*Bist du ein Fan von Bayern München?*“ werden demnach durch Ausführung des Tagging-Prozesses zunächst die Konzeptterme „Fan“ und „Bayern München“ identifiziert. Im zweiten Schritt wird durch Anwendung der Abbildungsfunktion 5.1 für jeden dieser Konzeptterme eine gewichtete Liste an Wikipedia-Artikeln ermittelt, von denen die ersten zehn Artikel als mögliche Konzepte des Dialogbeitrags definiert werden.

Kategorisierung

Nachdem die angesprochenen Konzepte bekannt sind und in Form von Wikipedia-Artikeln repräsentiert werden, ist die Spezifikation ihrer thematischen Zuordnungen und somit die Identifikation der in der Äußerung enthaltenen Themenvorschläge erforderlich. Dabei ist zu beachten, dass ein Konzept mit mehr als nur einem Thema in Beziehung stehen kann, wenn auch in unterschiedlichem Ausmaß. Das Konzept „Fan“ beispielsweise lässt sich den Themen „Kultur“, „Sport“ und „Gesellschaft“ zuordnen. Die Höhe der Relevanz dieser Themen für das Konzept lässt sich in Form von **Gewichten** ausdrücken. Durch Definition eines **Vektorraums**, dessen Raumdimensionen je

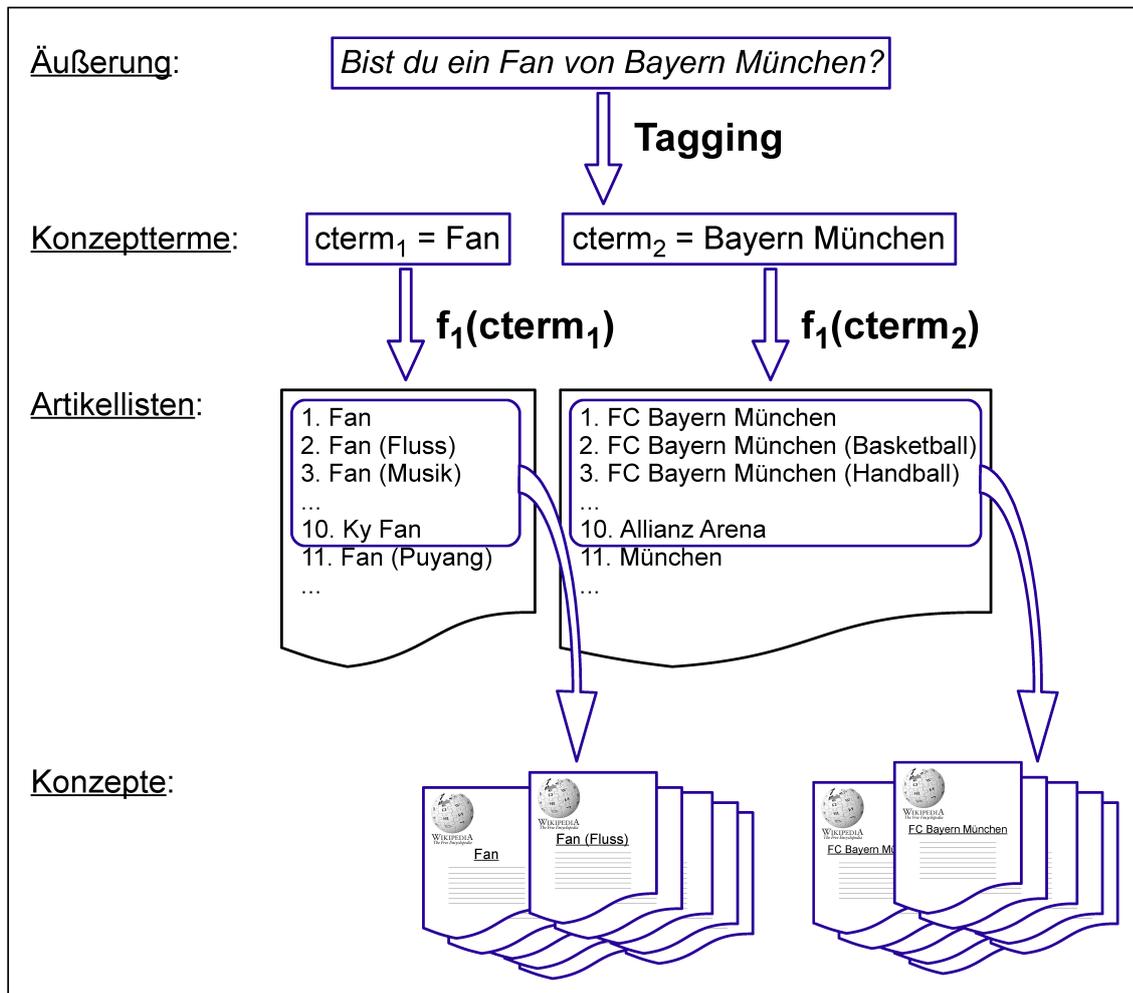


Abbildung 5.6: Konzeptualisierung der Äußerung „Bist du ein Fan von Bayern München?“.

einem möglichen Thema des betrachteten Dialogs entsprechen, lässt sich ein Konzept daher als Vektor innerhalb dieses Raums darstellen (s. *Vektorraummodell*, Abschnitt 3.2.1). Abbildung 5.7 veranschaulicht die Vektordarstellung des Konzepts „Fan“ auf Basis der genannten Themen „Kultur“, „Sport“ und „Gesellschaft“. Diese Darstellung liefert die Grundlage für die Bestimmung der **Ähnlichkeit** zwischen zwei Konzepten hinsichtlich ihrer Themenzugehörigkeit und darüber hinaus die Basis für die im nachfolgenden Abschnitt beschriebene Identifikation thematischer Überschneidungen zwischen zwei Dialogbeiträgen. Voraussetzung dafür ist jedoch die Definition einer festen Menge an möglichen Themen.

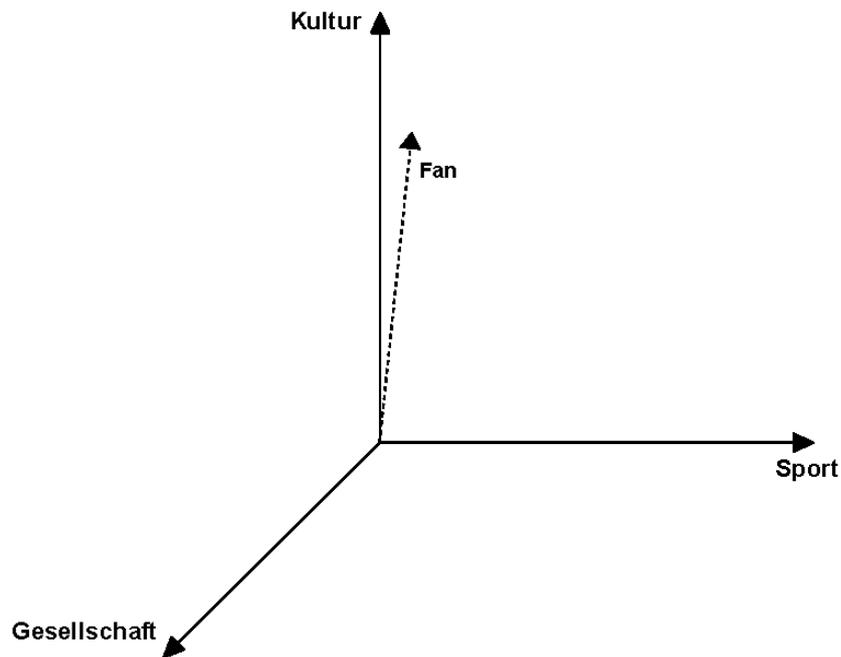


Abbildung 5.7: Vektor-basierte Darstellung der Relevanz der Dialogthemen „Sport“, „Kultur“ und „Gesellschaft“ für das Konzept „Fan“.

Innerhalb des Themenerkenners wird die Ermittlung der Themengewichte der einzelnen Konzepte durch eine Abbildung der diese Konzepte repräsentierenden Artikel auf eine Menge an vordefinierten Wikipedia-Kategorien umgesetzt. Zu diesem Zweck wird im Vorfeld eine Anzahl an bestimmten Kategorien festgelegt, die die Menge möglicher Themen für eine gegebene Dialogsituation wiedergeben. Dafür kommen prinzipiell alle in der Wikipedia enthaltenen Kategorien in Frage. Es bietet sich jedoch an, solche Kategorien als potentielle Dialogthemen festzulegen, die einen hohen Abstraktionsgrad aufweisen und somit allgemeinere Themengebiete abdecken, wie beispielsweise „Sport“ oder „Politik“, um die Menge klein und überschaubar zu halten. Je nach Dialogsituation kann es allerdings auch sinnvoll sein, spezifischere Kategorien und somit Dialogthemen zu bestimmen.

Ein Vorteil der Eingrenzung der Kategorienmenge auf eine überschaubare Anzahl an möglichen Themen ist die in Abschnitt 3.2.2 erläuterte Tatsache, dass die von Wikipedia bereitgestellte Menge an Kategorien zu umfassend ist. Sie enthält demnach eine große Anzahl an sehr spezifischen Titeln, wie „Verein im Fußball- und Leichtathletik-Verband Westfalen“, sowie sehr generelle wie „Frau“, die sich nicht für die Bezeichnung

von Dialogthemen eignen. Darüber hinaus geht eine Eingrenzung der Themenmenge bei einer Integration des aus der automatischen Themenerkennung hervorgehenden Informationen in ein regelbasiertes Dialogsystem mit einem geringeren Aufwand bezüglich der vorzunehmenden Regeländerungen und -anpassungen einher. Dieser Aspekt wird im nachfolgenden Kapitel 6 ausführlicher diskutiert. Um mit der Eingrenzung der Themenmenge weitere Kosten zu sparen, enthält der Kategorienindex (s. Abschnitt 5.1.2) lediglich Dokumente für die zuvor spezifizierten Kategorien. Dadurch wird nicht nur der Speicheraufwand reduziert, sondern auch ein schnellerer Zugriff auf die im Index abgelegten Informationen gewährleistet.

Die Gewichtung der Konzeptthemen geschieht durch die Abbildung der im Rahmen der Konzeptualisierung erstellten Menge an Konzepten C_{utt} auf den Vektorraum V . Dessen Dimensionen entsprechen den Elementen der Menge aller vordefinierten Wikipedia-Kategorien Cat_{wiki} . Grundlage hierfür bildet eine zweite Abbildungsfunktion f_2 , welche jeder einzelnen Kategorie $cat \in Cat_{wiki}$ ein Gewicht w für die Relevanz dieser Kategorie für das jeweils betrachtete Konzept $con \in C_{utt}$ zuweist. Hierbei gilt zu beachten, dass jedes Konzept con einem Wikipedia-Artikel entspricht. Auf diese Weise lässt sich con als n -dimensionaler Vektor $\vec{v}_{con} = (w_1, w_2, \dots, w_n)$ in V darstellen, wobei n die Anzahl aller Elemente in Cat_{wiki} ist. Realisiert wird f_2 auf Basis des Kategorienindex (s. Abschnitt 5.1.2). con wird also auf die Menge aller Dokumente des Kategorienindex D_{cat} abgebildet, wobei jedes einzelne Dokument $dCat \in D_{cat}$ eine vordefinierte Kategorie $cat \in Cat_{wiki}$ repräsentiert:

$$f_2 : con \mapsto D_{cat} \text{ mit } \forall cat \in Cat_{wiki} \exists dCat \in D_{cat}$$

Um anhand von f_2 jedes Konzept con durch einen Vektor \vec{v}_{con} zu repräsentieren, wird jedem in dem Kategorienindex enthaltenen Dokument $dCat_i \in D_{cat}$ auch hier auf Grundlage der *Lucene*-internen Ähnlichkeitsfunktion $score(con, dCat)$ (s. Formel 5.2) ein Gewicht w_i zugeordnet. Somit wird die Kategorisierung anhand folgender Funktion realisiert:

$$\boxed{f_2(con) = \vec{v}_{con} = (w_1, w_2, \dots, w_n) \text{ mit } n = |Cat_{wiki}|} \quad (5.3)$$

Auf Grundlage dieser Abbildung lässt sich für jedes in der aktuell betrachteten Äußerung enthaltene, zuvor identifizierte Konzept ein Vektor bestimmen, dessen Dimen-

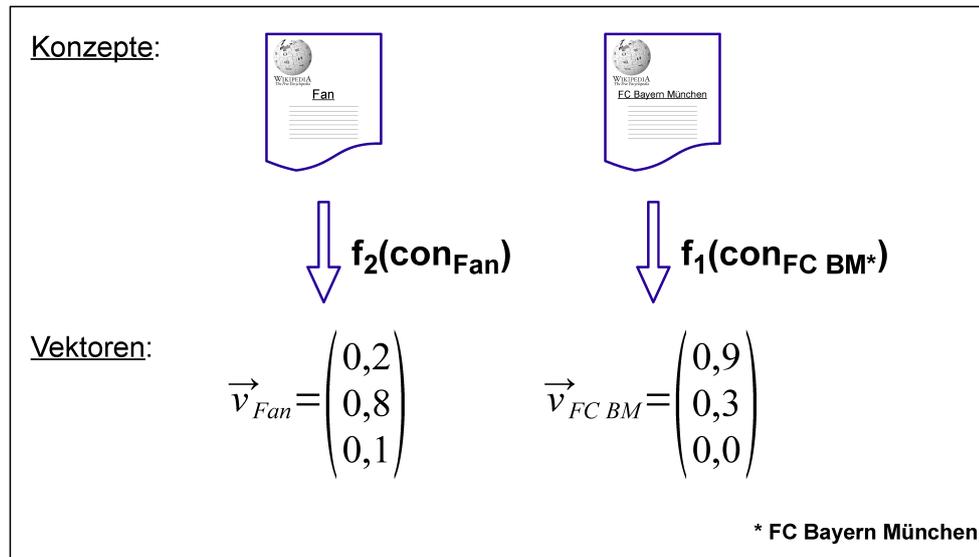


Abbildung 5.8: Kategorisierung der Konzepte „Fan“ und „FC Bayern München“.

sion n mit der Anzahl der vordefinierten Kategorien übereinstimmt. Jeder Eintrag dieses Vektors entspricht demnach einem möglichen Dialogthema und die einzelnen Werte dieser Einträge stellen den Grad der Relevanz des jeweiligen Themas für das betrachtete Konzept dar. Im weiteren Verlauf der vorliegenden Ausarbeitung werden diese im Rahmen der Erkennung von Themenvorschlägen ermittelten Vektoren **Konzeptthemen-Vektoren** genannt.

Beispiel Der Prozess der Kategorisierung lässt sich anhand des in Abbildung 5.8 dargestellten Beispiels zusammenfassen. Unter der Annahme, dass die durch die Abbildungsfunktion f_2 ermittelten Relevanzwerte zwischen 0 und 1 liegen, zuvor drei unterschiedliche Themen durch die Wikipedia-Kategorien „Sport“, „Kultur“ und „Gesellschaft“ definiert wurden (sodass $n = 3$) und diese den in Abbildung 5.9 dargestellten Vektorraum aufspannen, ließen sich für die beiden Konzepte „Fan“ und „FC Bayern München“ je ein Konzeptthemen-Vektor ermitteln. Anhand dieser Vektoren lassen sich die thematischen Zugehörigkeiten der Konzepte erkennen. So stellt „Kultur“ mit einem Wert von 0,8 das relevanteste Thema für das Konzept „Fan“ dar, das Konzept „FC Bayern München“ hingegen gehört mit einem Relevanzwert von 0,9 eindeutig dem Thema „Sport“ an.

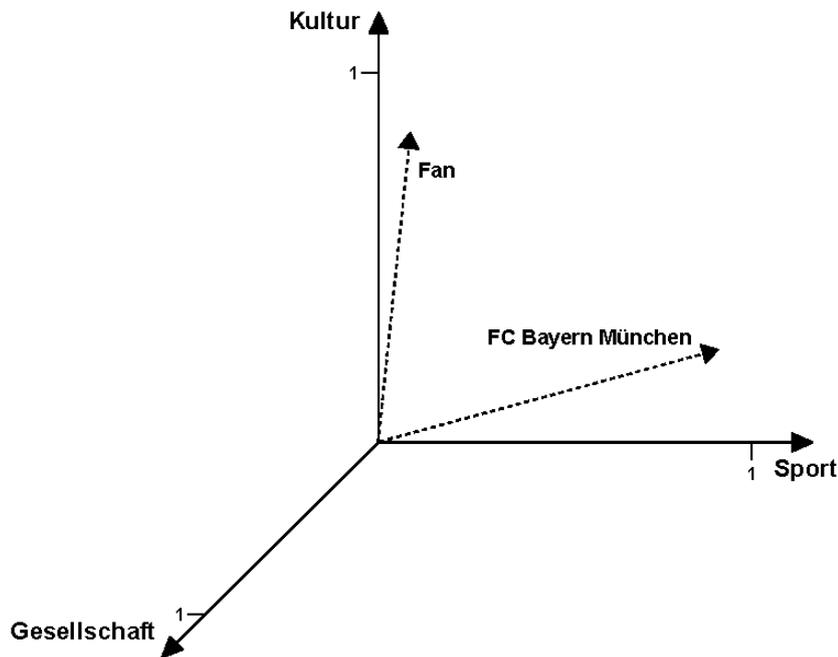


Abbildung 5.9: Vektor-basierte Darstellung der Relevanz der Dialogthemen „Sport“, „Kultur“ und „Gesellschaft“ für die Konzepte „Fan“ und „FC Bayern München“.

5.3.2 Identifikation des Dialogthemas

Mit dem Ziel, das aktuelle Thema des betrachteten Dialogs zu identifizieren, erfolgt im zweiten Basisprozess des Themenerkenners auf Grundlage der zuvor identifizierten Konzeptthemen-Vektoren die Identifikation des Themas, hinsichtlich dessen sich zwei aufeinander folgende Äußerungen überschneiden. Die **Berechnung thematischer Überschneidungen** zwischen zwei Dialogbeiträgen geschieht mittels eines Ähnlichkeitsvergleichs jeder einzelnen der für die erste Äußerung ermittelten Konzeptthemen-Vektoren mit jedem der für die anschließende Äußerung ermittelten Konzeptthemen-Vektoren. Realisiert wird dieser Vergleich durch Anwendung des in Abschnitt 3.2.1 eingeführten **Kosinus-Ähnlichkeitsmaß**. Die Ähnlichkeit zwischen zwei Konzepten $con1$ und $con2$ zwei aufeinander folgender Äußerungen $utt1$ und $utt2$ wird dabei auf Basis ihrer zuvor bestimmten Vektordarstellungen \vec{v}_{con1} und \vec{v}_{con2} mittels

$$sim(con1, con2) = \frac{\vec{v}_{con1} \bullet \vec{v}_{con2}}{|\vec{v}_{con1}| \cdot |\vec{v}_{con2}|} \quad (5.4)$$

berechnet wird, wobei $con1 \in C_{utt1}$ und $con2 \in C_{utt2}$. Dieser Ähnlichkeitsvergleich

wird nun für alle Konzeptpaare (con_{utt1}, con_{utt2}) des *kartesischen Produkts* der Konzeptmengen C_{utt1} und C_{utt2} durchgeführt, sodass die Ähnlichkeit zwischen diesen Konzeptmengen als die Menge aller Konzeptähnlichkeiten SIM zwischen $utt1$ und $utt2$

$$sim(C_{utt1}, C_{utt2}) := SIM_{utt1,utt2} = \{sim(a, b) \mid sim(a, b) : (a, b) \in C_{utt1} \times C_{utt2}\} \quad (5.5)$$

definiert ist. Lässt sich im Rahmen dieser Vergleiche eine Ähnlichkeit zwischen zwei Konzepten con_{utt1} und con_{utt2} der Äußerungen berechnen, deren Wert mindestens 0,5 ist, also $sim(con_{utt1}, con_{utt2}) \geq 0,5$, wird eine thematische Überschneidung zwischen diesen Konzepten und somit zwischen den Äußerungen $utt1$ und $utt2$ identifiziert. Der Wert spiegelt dabei die *Sicherheit* für einen thematischen Zusammenhang zwischen diesen Äußerungen wieder, die aus subjektiver Sicht ab 50% ausreichend für die Identifikation einer thematischen Überschneidung ist.

Hinsichtlich welchen Themas sich zwei Äußerungen überschneiden, lässt sich anschließend anhand des sogenannten **Dialogthemen-Vektors** bestimmen, dessen einzelne Einträge die Relevanz jeder vordefinierten Wikipedia-Kategorie und somit jedes möglichen Dialogthemas für die beiden betrachteten Dialogbeiträge definiert. Zu diesem Zweck werden zunächst alle Konzeptthemen-Vektoren, die an einer thematischen Überschneidung zwischen utt_1 und utt_2 beteiligt sind, in einer Menge $CV_{topOverlap}$ vereint. Die Vektoren der Konzeptterme, die an mehreren dieser Überschneidungen beteiligt sind, werden dabei entsprechend häufig hinzugefügt, um den Einfluss der jeweiligen Konzepte auf das Dialogthema zu maximieren.

Der **Dialogthemen-Vektor** errechnet sich dann aus der Summe aller in $CV_{topOverlap}$ enthaltenen Konzeptthemen-Vektoren, sodass

$$\vec{v}_{dialog} = \sum_{i=1}^n \vec{v}_i \text{ mit } \vec{v}_i \in CV_{topOverlap}, n = |CV_{topOverlap}| \quad (5.6)$$

Abschließend werden die einzelnen Einträge des resultierenden Vektors durch die Anzahl der aufsummierten Konzeptthemen-Vektoren normiert, um Relevanzwerte für die einzelnen Kategorien zu erhalten, die zwischen 0 und 1 liegen. Sobald einer dieser Werte einen bestimmten, zuvor festzulegenden Schwellwert übersteigt, wird die zu diesem Wert gehörende Kategorie als **aktuelles Dialogthema** identifiziert.

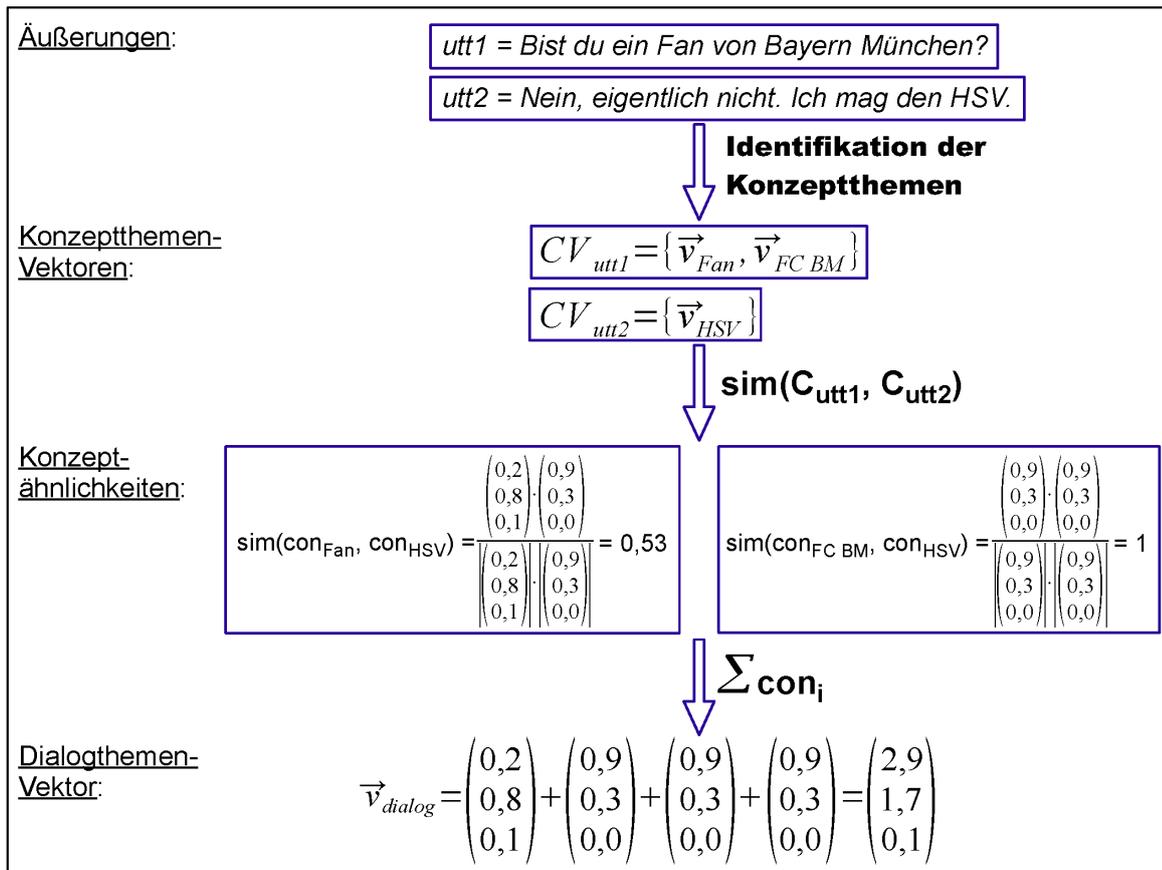


Abbildung 5.10: Berechnung des Dialogthemenvektors auf Basis der Konzeptthemenvektoren von zwei Äußerungen.

Beispiel Der Prozess zur Identifikation des aktuellen Dialogthemas lässt sich anhand des in Abbildung 5.10 skizzierten Beispiels demonstrieren. Voraussetzung für die Identifikation thematischer Überschneidungen zwischen den zwei Äußerungen *utt1* und *utt2* ist, dass für beide aufeinander folgende Beiträge der Prozess der Konzeptthemen-Identifikation (s. vorheriger Abschnitt) durchgeführt wurde und dem System somit die Konzeptthemen-Vektoren der beiden Beiträge vorliegen. Mit dem Ziel, diese beiden Äußerungen anschließend auf thematische Überschneidungen zu überprüfen, wird $\text{sim}(\text{con}_1, \text{con}_2)$ (Gleichung 5.4) zweimal durchgeführt, einmal zur Überprüfung der Ähnlichkeit zwischen den Konzepten „Fan“ und „Hamburger SV“ und einmal zur Überprüfung der Ähnlichkeit zwischen den Konzepten „FC Bayern München“ und „Hamburger SV“. Beide Vergleiche ergeben eine Kosinus-Ähnlichkeit $> 0,5$, sodass alle an den Vergleichen beteiligten Vektoren zu einem Vektor addiert werden. Der Vektor für

das Konzept „FC Bayern München“ wird dabei zweifach addiert, da dieser an zwei thematischen Überlappungen beteiligt ist.

Der resultierende Dialogthemen-Vektor liefert nun die Grundlage für die Identifikation des gemeinsamen Themas der beiden Äußerungen und somit des aktuellen Dialogthemas. Die einzelnen Einträge des Vektors werden dafür durch die Anzahl der Summanden, das heißt durch die Anzahl der aufsummierten Konzeptthemen-Vektoren, normiert, um Relevanzwerte für die einzelnen Kategorien zu erhalten, die zwischen 0 und 1 liegen. Sobald einer dieser Werte einen bestimmten Schwellwert übersteigt (die Festlegung dieses Schwellwerts erfolgt in Abschnitt 5.4.2), wird die zu diesem Wert gehörende Kategorie als aktuelles Dialogthema identifiziert. Für die betrachteten Äußerungen lassen sich folgende Relevanzwerte zusammentragen:

Sport:	2,9 : 4 =	0,725
Kultur:	1,7 : 4 =	0,425
Gesellschaft:	0,1 : 4 =	0,025

Unter der Annahme, dass der zuvor festgelegte Schwellwert auch hier bei 0,5 liegt, wird die Kategorie „Sport“ als Dialogthema der exemplarisch skizzierten Dialogsequenz identifiziert. Für den Fall, dass die beschriebenen Konditionen zur Identifikation eines Dialogthemas innerhalb eines Prozesses von mehreren Kategorien gleichzeitig erfüllt werden, liegt kein eindeutiges Ergebnis vor, sodass zu dem Zeitpunkt mehrere Optionen für das aktuelle Dialogthema existieren.

5.3.3 Algorithmus

Die technische Umsetzung der Basisprozesse zur automatischen Erkennung des aktuellen Dialogthemas erfolgt auf Basis des Algorithmus 1 (s. Seite 116). Der erste Basisprozess, die *Identifikation der Konzeptthemen* (s. Zeile 2–3), wird somit für jede Äußerung des betrachteten Dialogs ausgeführt, der zweite Basisprozess, die *Identifikation des Dialogthemas* (s. Zeile 5–7), wird hingegen erst mit der zweiten Äußerung des Dialogs für jede weitere Äußerung wiederholt (s. Zeile 5). Die einzelnen Schritte dieser Prozesse werden mit Hilfe von Methoden und den entsprechenden, zuvor hergeleiteten Funktionen realisiert (ab Zeile 12). Die Ausführung dieses Algorithmus resultiert in der stetigen Aktualisierung des Dialogthemas mit jeder neu verarbeiteten Äußerung. Das Dialogthema entspricht dabei der Kategorie, deren Koordinate, durch die sie innerhalb

des Dialogthemen-Vektors repräsentiert wird, für die jeweils zwei betrachteten Äußerungen einen bestimmten Schwellwert erreicht (s. Zeile 33 – 34). Gilt dies für mehrere Kategorien, so existieren mehrere Dialogthemen zu einem Zeitpunkt.

Nicht immer führt die Ausführung des Algorithmus zu dem gewünschten Ergebnis, zum Beispiel, wenn die Dialogteilnehmer das Thema wechseln. Mit dem Ziel, solche und weitere, den thematischen Verlauf des Dialogs beeinflussenden Aktionen automatisch zu erkennen, führt der Themenerkenners simultan zu den Basisprozessen weitere Aufgaben durch, die die Ableitung entsprechender Themeninformationen ermöglichen. Abschnitt 5.3.4 liefert eine Beschreibung dieser Aufgaben.

Graphische Darstellung der Ergebnisse

Um den Benutzern einen Einblick in die mit jedem Durchlauf des Algorithmus ermittelten Informationen zu gewähren, wurde eine graphische Darstellung der Ergebnisse des Themenerkenners realisiert. In Abbildung 5.11 ist ein Beispiel dieser Darstellung, bestehend aus einer *Dialogoberfläche* und einer *Themenoberfläche*, zu sehen. Die Dialogoberfläche, hier als Oberfläche ① markiert, liefert eine Übersicht über die aktuell verarbeiteten Äußerungen des betrachteten Dialogs. Neben den einzelnen Äußerungen werden das jeweils identifizierte Dialogthema mit zugehörigem Relevanzwert und die Information darüber, ob ein möglicher Themensprung vorliegt, angezeigt.

Die Themenoberfläche, hier als Oberfläche ② markiert, gibt einen Überblick über die aktuellen Relevanzwerte für die einzelnen, die möglichen Dialogthemen repräsentierenden Kategorien in Form eines Balkendiagramms. Für jede Kategorie ist ein Balken in einer anderen Farbe vorhanden. Zu Gunsten der Übersichtlichkeit werden dabei nur diejenigen Balken angezeigt, deren Kategorien einen Wert > 0 aufweisen. Die Höhe der Balken spiegelt die Relevanz der Kategorien und somit die Wahrscheinlichkeit (engl. *probability*) des dadurch repräsentierten Dialogthemas für die aktuelle Dialogsituation wieder. Die gestrichelte Linie verdeutlicht die Höhe des aktuell eingestellten Schwellwerts, der hier bei 0,3 liegt.

Jede dieser Oberflächen wird mit jedem Durchlauf des Themenerkenners-Algorithmus, also mit der Verarbeitung einer weiteren Äußerung, aktualisiert. Der Benutzer kann den Dialogverlauf auf diese Weise besser nachvollziehen und die Dialogbeiträge ermitteln, die die Durchführung thematischer Aktionen ausgelöst haben. So liefern diese Oberflächen auch nachfolgend die Grundlage für die Darstellung einzelner Beispiele.

Algorithmus 1 Pseudocode: Automatische Themenerkennung

```

1: repeat
2:   for all {utterance ∈ dialog} do
3:      $C_{utt} \leftarrow \text{CONZEPTUALIZE}(utterance)$ 
4:      $CV_{utt} \leftarrow \text{CATEGORIZE}(C_{utt})$ 
5:     if  $CV_{previousUtt} \neq 0$  then
6:        $CV_{topOverlap} \leftarrow \text{GETTOPICALOVERLAPS}(CV_{utt}, CV_{previousUtt})$ 
7:        $dialogTopic \leftarrow \text{GETDIALOGTOPIC}(CV_{topOverlap})$ 
8:     end if
9:      $CV_{previousUtt} \leftarrow CV_{utt}$ 
10:  end for
11: until end of dialog

12:  $\text{CONZEPTUALIZE}(utterance)$ 
13: for all {cterm ∈ utterance} do
14:    $C_{utt}.add(f_1(cterm))$ 
15: end for
16: return  $C_{utt}$ 

17:  $\text{CATEGORIZE}(C_{utt})$ 
18: for all {con ∈  $C_{utt}$ } do
19:    $CV_{utt}.add(f_2(con))$ 
20: end for
21: return  $CV_{utt}$ 

22:  $\text{GETTOPICALOVERLAPS}(CV_{utt}, CV_{previousUtt})$ 
23: for all { $sim(con_i, con_j) \in SIM_{utt, previousUtt}$ } do
24:   if { $sim(con_i, con_j) \geq 0.5$ } then
25:      $CV_{topOverlap}.add(f_2(con_i), f_2(con_j))$ 
26:   end if
27: end for
28: return  $CV_{topOverlap}$ 

29:  $\text{GETDIALOGTOPIC}(CV_{topOverlap})$ 
30: for all { $\vec{v} \in CV_{topOverlap}$ } do
31:    $\vec{v}_{dialog} \leftarrow \vec{v}_{dialog} + \vec{v}$ 
32: end for
33: for all { $coordinate(category) \in \vec{v} > threshold$ } do
34:   return category
35: end for

```

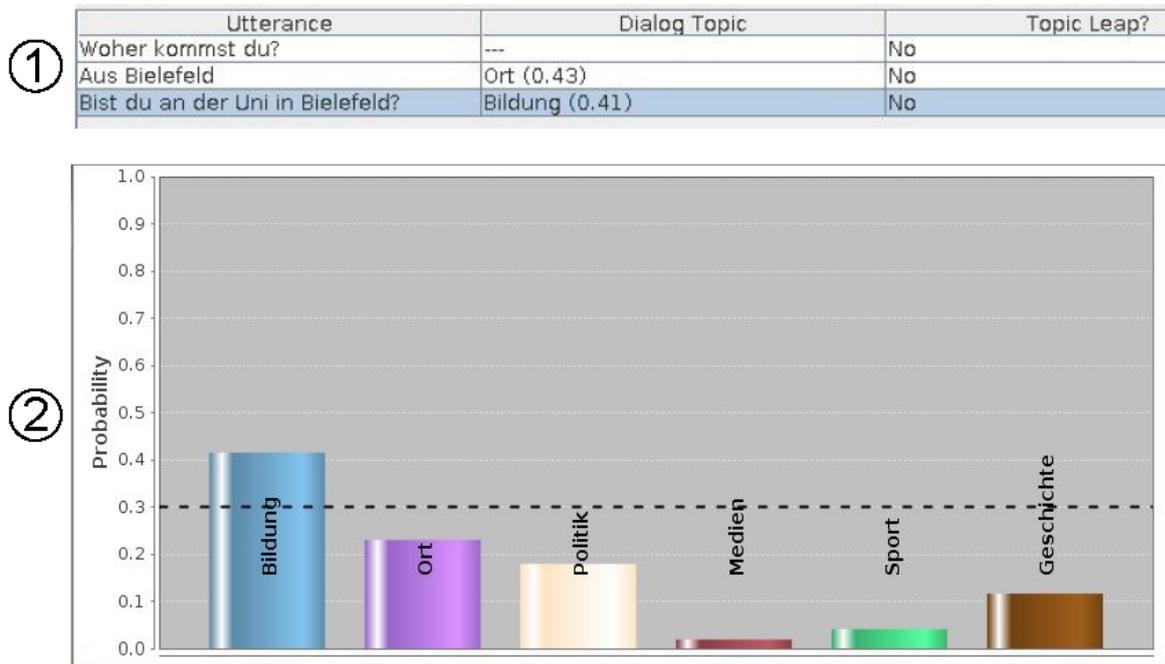


Abbildung 5.11: Graphische Darstellung der aktuellen Ergebnisse des Themenerkenners. Oberfläche ① liefert eine Übersicht über die bereits verarbeiteten Äußerungen. Oberfläche ② gibt einen Überblick über die aktuellen Relevanzwerte für die einzelnen, die möglichen Dialogthemen repräsentierenden Kategorien.

5.3.4 Ableitung der Themeninformationen

Wie zu Beginn dieses Kapitels festgehalten, sind simultan zu der Durchführung der beschriebenen Basisprozesse die unterschiedlichen, in Abschnitt 4.3.3 definierten Aufgaben zur Identifikation von Informationen über die aktuelle Dialogsituation auszuführen. Diese erfordern keine Umsetzung weiterer Prozesse, sondern lassen sich, wie nachfolgend erläutert, nebenläufig realisieren.

Aufgabe 1: Identifikation von Themenvorschlägen

Mit der *Identifikation der Konzeptthemen* werden implizit Informationen hinsichtlich der mit einem Dialogbeitrag eingeführten thematischen Richtungen und somit **mögliche Dialogthemen** ermittelt. Die Höhe der Wahrscheinlichkeit eines einzelnen Themas errechnet sich dabei aus der Summe aller Konzeptthemen-Vektoren CV der jeweiligen Äußerung. Die *Identifikation von Themenvorschlägen* resultiert daher in einem **Beitragsthemen-Vektor**

$$\vec{v}_{utt} = \sum_{i=1}^n \vec{v}_i \text{ mit } \vec{v}_i \in CV_{utt}, n = |CV_{utt}| \quad (5.7)$$

Die einzelnen Koordinaten des Beitragsthemen-Vektors repräsentieren die möglichen Themen einer Äußerung, ihre Werte geben Aufschluss über ihre Wahrscheinlichkeiten. Das heißt, jede Koordinate mit einem Wert > 0 stellt einen Themenvorschlag und somit ein mögliches Dialogthema dar und wird in der Information über eingeführte Themen erfasst. Ob sich letztendlich eines dieser Themen als Dialogthema durchsetzt und wenn ja, welches, hängt von der darauf folgenden Äußerung des anderen Dialogteilnehmers und den darin angesprochenen Konzepten ab.

Abgesehen von der Identifikation etwaiger Vorschläge für Dialogthemen, erlaubt der Beitragsthemen-Vektor die Spezifikation eines *Satzthemas* (s. Abschnitt 2.2.1), das heißt des vorherrschenden Themas der aktuell betrachteten Äußerung. Hierfür wird die Koordinate mit dem höchsten Wert in \vec{v}_{utt} ermittelt. Bezüglich der Beispieläußerung „*Bist du ein Fan von Bayern München?*“ lässt sich das **Satzthema** wie folgt spezifizieren:

$$\vec{v}_{uttExp} = \vec{v}_{Fan} + \vec{v}_{FCBM} = \begin{pmatrix} 0, 2 \\ 0, 8 \\ 0, 1 \end{pmatrix} + \begin{pmatrix} 0, 9 \\ 0, 3 \\ 0, 0 \end{pmatrix} = \begin{pmatrix} 1, 1 \\ 1, 1 \\ 0, 1 \end{pmatrix}$$

Da im vorliegenden Fall zwei Koordinaten den höchsten Wert aufweisen, werden der Äußerung entsprechend dazu zwei Satzthemen zugeordnet, nämlich „Sport“ und „Kultur“. Das auf diese Weise identifizierte Satzthema dient der Generierung einer entsprechenden Reaktion des Agenten Max auf die von seinem menschlichen Dialogpartner initiierten Themeneinführungen. Nähere Informationen hierzu liefert Abschnitt 6.3.2.

Aufgabe 2: Identifikation eines Dialogthemas

Mit Durchführung des Basisprozesses zur *Identifikation des Dialogthemas* wird gleichzeitig die zweite Aufgabe des Themenerkenners bewältigt. Eine Themenfortführung wird demnach mit jeder erfolgreichen Durchführung dieses Prozesses erkannt und resultiert in der Erstellung, beziehungsweise Aktualisierung der Information über das identifizierte Dialogthema. Ein Beispiel für eine erfolgreiche Identifikation eines Dialogthemas durch Fortführung eines mit der ersten Äußerung eines Dialogs vorgeschla-

genen Themas liefert Abbildung 5.12.

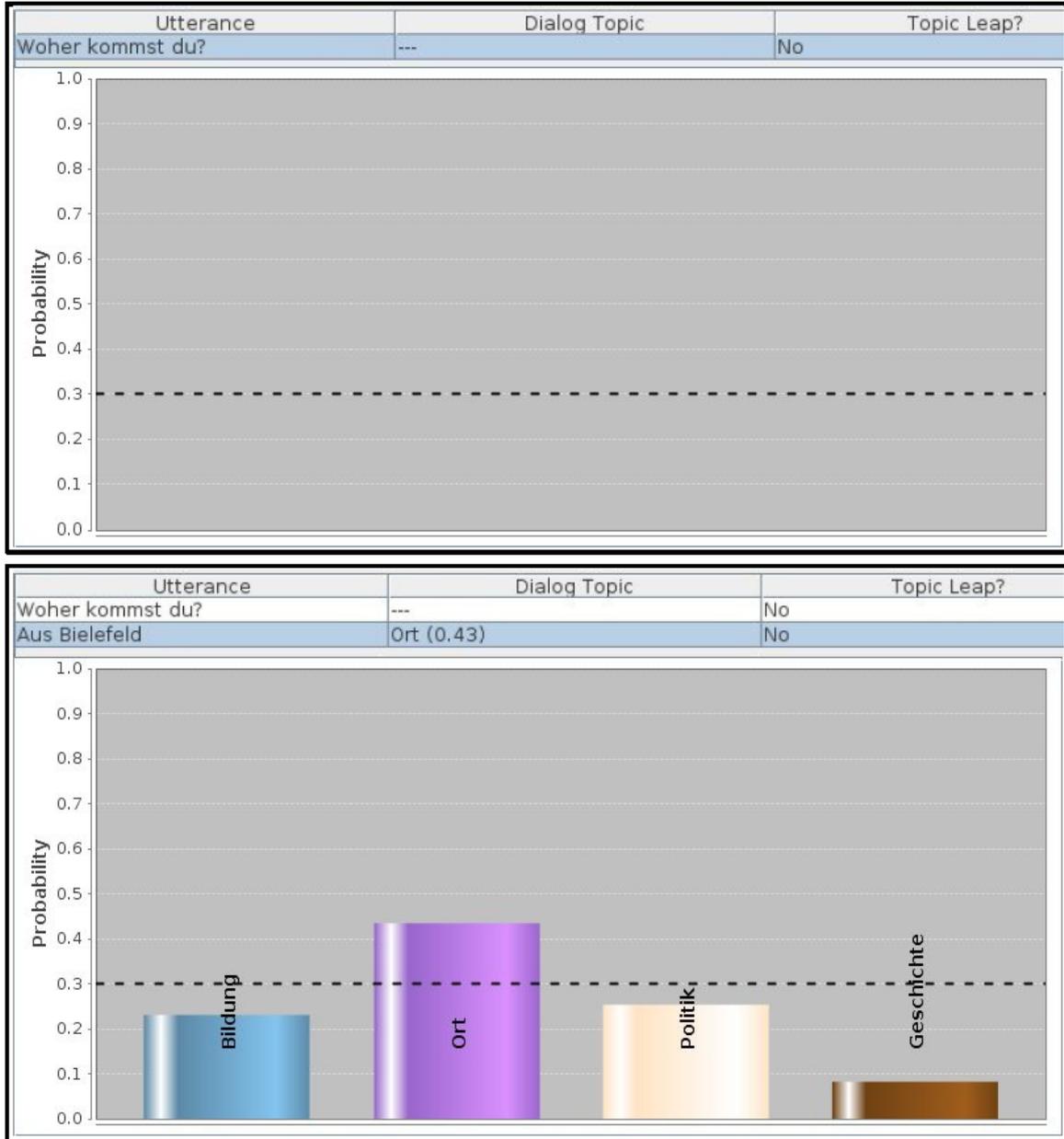


Abbildung 5.12: Identifikation des Dialogthemas „Ort“ auf Basis der ersten zwei Äußerungen eines Dialogs.

Lässt sich allerdings keine thematische Überschneidung der in der aktuellen Äußerung enthaltenen Konzepte zu den Konzepten der vorangegangenen Äußerung identifizieren, wird keine aktuelle Themeninformation hinsichtlich des Dialogthemas generiert.

Liegt bis zu diesem Zeitpunkt keine aktuelle Information vor, konnte demnach für den gesamten bisherigen Dialogverlauf und auch mit der aktuell betrachteten Äußerung noch kein Dialogthema bestimmt werden. Existiert allerdings eine entsprechende Information, sodass vor der aktuellen Dialogsituation bereits ein Dialogthema identifiziert wurde, deutet das Fehlschlagen der Ermittlung thematischer Überschneidungen auf Basis der aktuellen Äußerung auf einen möglichen *Wechsel* des aktuellen Dialogthemas hin.

Aufgabe 3: Erkennung eines möglichen Themensprungs

Führt eine Äußerung das zuvor identifizierte Dialogthema nicht weiter fort und weist keinerlei konzeptionelle Verbindung zu diesem Dialogthema auf, kann sich ein *Themensprung* ereignen. Ein möglicher Themensprung kündigt sich also an, wenn sich im Rahmen des Vergleichs der Konzepte der ersten Äußerung *utt1* mit der zweiten Äußerung *utt2* anhand $\text{sim}(C_{utt1}, C_{utt2})$ keine thematische Überschneidung > 0 ermitteln lässt. Darüber hinaus muss die das aktuelle Dialogthema repräsentierende Koordinate $\text{coordinate}(\text{dialogTopic})$ innerhalb des aktuellen Konzeptthemen-Vektors \vec{v}_{dialog} einen Wert $= 0$ aufweisen. Die Aktualisierung der entsprechenden Themeninformation erfolgt daher innerhalb der GETTOPICALOVERLAPS-Methode (s. Algorithmus 2, Zeile 5–8). Abbildung 5.13 liefert ein Beispiel für einen erkannten und mit der darauffolgenden Äußerung letztendlich durchgeführten Themensprung.

Algorithmus 2 Pseudocode: Erkennung eines möglichen Themensprungs

```

1: GETTOPICALOVERLAPS( $CV_{utt}, CV_{previousUtt}$ )
2: for all  $\{\text{sim}(\text{con}_i, \text{con}_j) \in \text{SIM}_{utt,previousUtt}\}$  do
3:   if  $\{\text{sim}(\text{con}_i, \text{con}_j) \geq 0.5\}$  then
4:      $CV_{topOverlap}.\text{add}(f_2(\text{con}_i), f_2(\text{con}_j))$ 
5:      $\text{topicLeap} \leftarrow \text{NO}$ 
6:   else
7:     if  $\{\text{coordinate}(\text{dialogTopic}) \in \vec{v}_{con} = 0\}$  then
8:        $\text{topicLeap} \leftarrow \text{YES}$ 
9:     end if
10:  end if
11: end for

```

5.3 Wikipedia-basierte Umsetzung des Themenerkenners

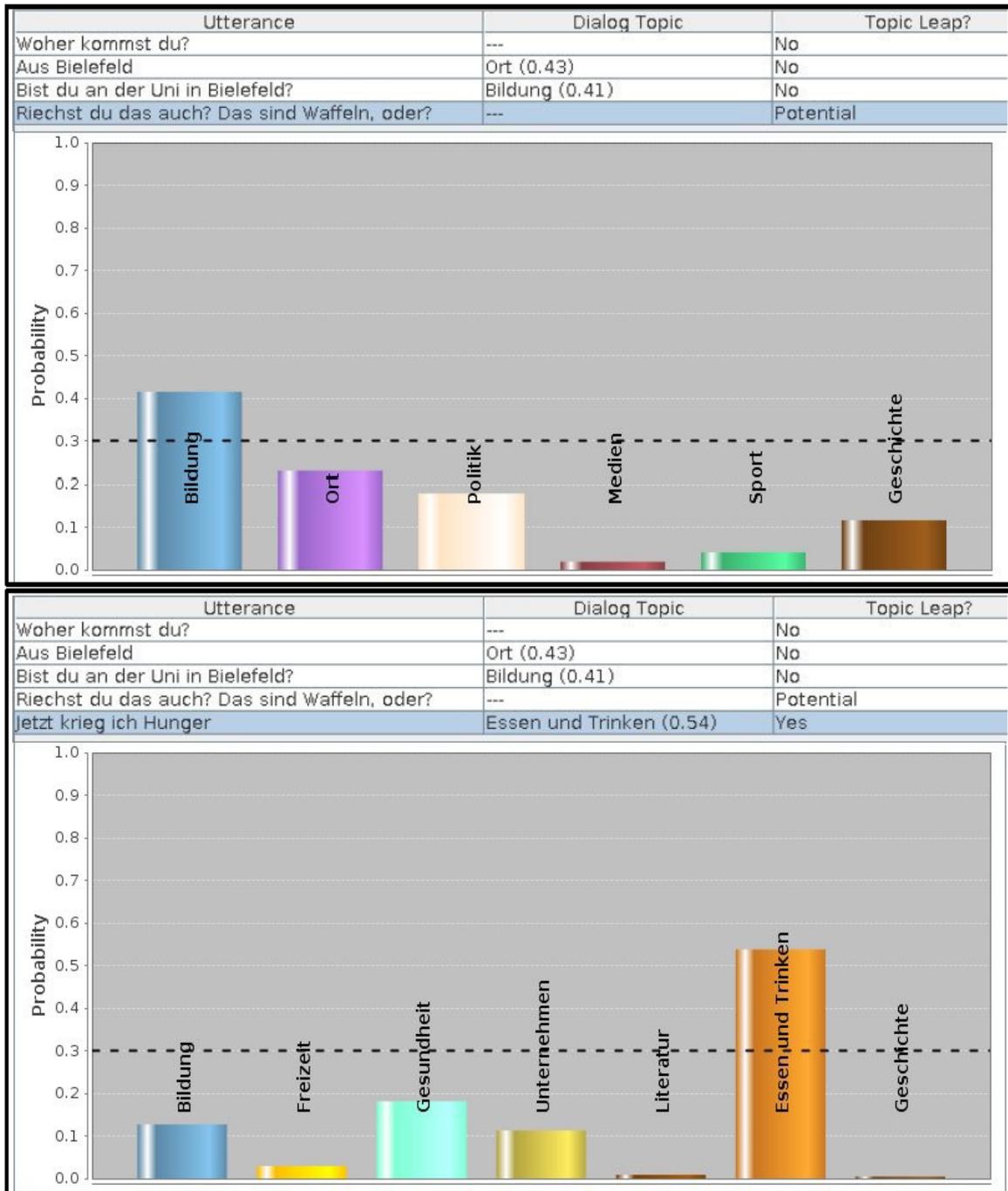


Abbildung 5.13: Erkennung eines möglichen und durchgeführten Themensprungs.

Aufgabe 4: Identifikation thematischer Überschneidungen zu früheren Äußerungen

Mit dem Ziel, den aktuellen thematischen Kontext eines Dialogs in die nachfolgenden Prozesse zur Identifikation des Dialogthemas mit einzubeziehen, wird bei den weiteren Identifikationsprozessen zusätzlich zu dem Vergleich der Konzeptthemen-Vektoren $sim(C_{utt1}, C_{utt2})$ ein Vergleich der Konzeptthemen-Vektoren der aktuellen Äußerung mit dem letzten Dialogthemen-Vektor durchgeführt, also $sim(C_{utt1}, \vec{v}_{dialog})$. Diese Art der Integration des thematischen Kontexts lässt sich mit beliebig vielen vergangenen Dialogthemen-Vektoren durchführen, sodass durch Dialogthemen-Vektoren repräsentiertes Wissen über zuvor angesprochene Themen die aktuelle Themenidentifikation beeinflusst und die Rückkehr zu einem vergangenen Dialogthema erleichtert wird. Auf diese Weise wird garantiert, dass thematisch irrelevante Zwischenbemerkungen keine große Auswirkung auf die Ermittlung des aktuellen Dialogthemas haben. Ein Beispiel für eine gelungene Themenwiedereinführung liefert Abbildung 5.14.

Aufgabe 5: Festlegung einer Themenbezeichnung

Wikipedia stellt Themenbezeichnungen in Form von eindeutigen Titeln für die einzelnen Kategorien zur Verfügung. Dementsprechend liegt der Ansatz nahe, ein Dialogthema mit dem Titel der Kategorie zu benennen, welche das jeweilige Dialogthema repräsentiert. Auf diese Weise müssen die einzelnen Themenbezeichnungen nicht ausdrücklich während des jeweiligen Dialogs genannt werden, da sie bereits im Vorfeld existieren. Wie bereits diskutiert, eignet sich dabei jedoch nur eine kleine Auswahl aller Kategorientitel als Bezeichnungen für Dialogthemen. Die bereits in Abschnitt 5.3.1 beschriebene Eingrenzung der Kategorienmenge auf eine überschaubare Anzahl an vordefinierten Kategorien ist ein erster Schritt, diese Problematik anzugehen. Darüber hinaus müssen einige Titel der verbleibenden Kategorien so abgewandelt werden, dass sie intuitiver klingen. Der Kategorientitel „Verkehrswesen“ zum Beispiel stellt eine eher unnatürliche Bezeichnung für ein Dialogthema dar. Hier muss daher eine Umbenennung durchgeführt werden. Der Titel wurde daher von Hand in die natürlicher klingende Themenbezeichnung „Verkehr“ geändert.

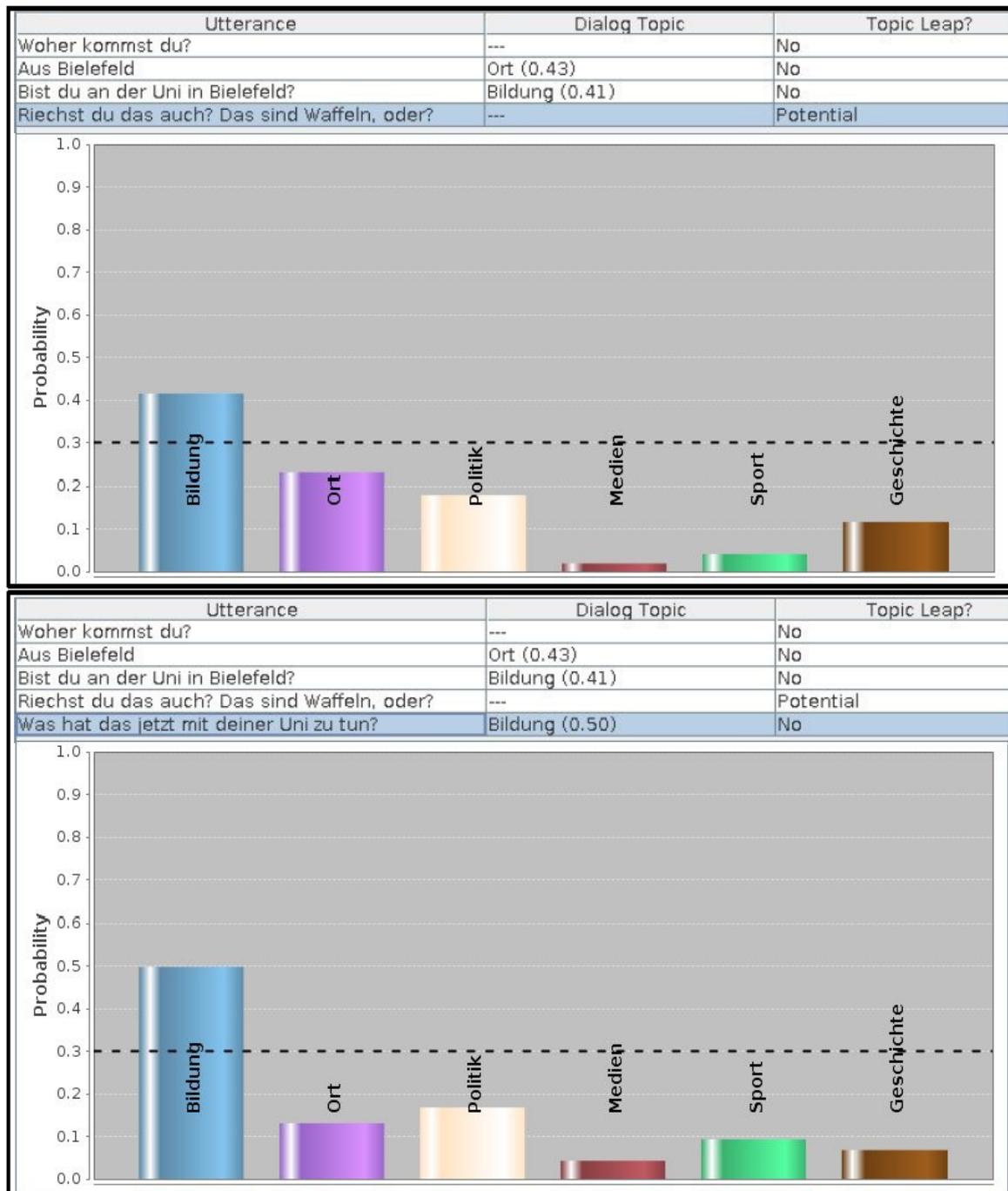


Abbildung 5.14: Beispiel für eine nach der Ermittlung eines möglichen Themensprungs vollzogene Themenwiedereinführung.

5.4 Evaluation

Mit dem Ziel, die Korrektheit des auf Grundlage des in diesem Kapitel umfassend vorgestellten Konzepts entwickelten Themenerkenners zu beurteilen, wurden insgesamt drei Evaluationen auf Basis von zwei unterschiedlichen Korpora durchgeführt. Bei der ersten handelte es sich dabei um eine vorbereitende Studie, die den grundlegenden Ansatz der automatischen Themenidentifikation und -benennung auf Basis kollaborativ erstellten Wissens in Form von Wikipedia-Informationen beurteilen sollte. Grundlage für diese Evaluation bildete eine Menge an thematisch klassifizierten Zeitungsartikeln. Die hier erzielten Ergebnisse verdeutlichten die grundsätzliche Tauglichkeit des Ansatzes, wiesen jedoch auch auf die Relevanz der Satz-übergreifenden Themenerkennung hin. Die zweite, ebenfalls auf den Zeitungsartikeln basierend durchgeführte Evaluation zeigte schließlich die Leistungsfähigkeit des gesamten, auf die Erkennung von thematischen Überschneidungen zwischen einzelnen Sätzen aufbauenden Ansatzes auf. Darüber hinaus gelang die Festlegung des Schwellwerts zur Identifikation des Dialogthemas. Im Anschluss daran erfolgte schließlich eine dritte Evaluation, welche Aufschluss über den Einsatz des gesamten Erkennungssystems in natürlichsprachlichen Dialogen und somit der eigentlichen Zielanwendung des innerhalb dieses Kapitels beschriebenen Konzepts gab. Nachfolgend werden alle drei Evaluationen und deren einzelne Ergebnisse näher vorgestellt. Zuvor sollen jedoch erst die bei der Evaluation zur Anwendung gekommenen Bewertungsmaße eingeführt werden.

5.4.1 Bewertungsmaße

Im Fachbereich des Information Retrieval (s. Abschnitt 3.2.1) kommen typischerweise zwei Maße zur Bewertung der Ergebnisse zum Einsatz: **Precision** und **Recall**. Während Precision die Fähigkeit, nur relevante Dokumente nachzuweisen, und somit die *Genauigkeit* eines Systems bestimmt, wird mit Recall die Fähigkeit, alle relevanten Dokumente nachzuweisen, und somit die *Vollständigkeit* eines Systems bewertet (Poetzsch, 2006). Precision und Recall lassen sich demnach anhand folgender Formeln berechnen:

$$Precision = \frac{\text{Anzahl der relevanten Dokumente in der Ergebnismenge}}{\text{Anzahl aller Dokumente in der Ergebnismenge}} \quad (5.8)$$

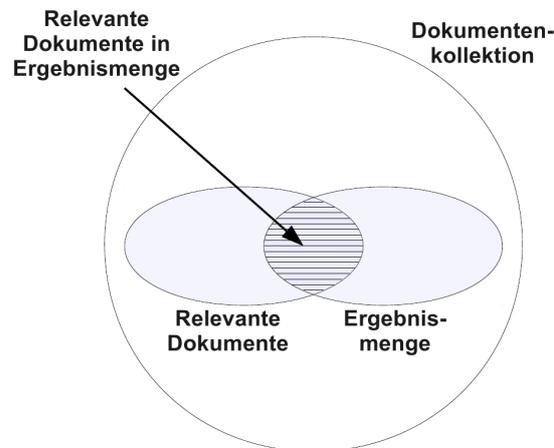


Abbildung 5.15: Precision (Genauigkeit) und Recall (Vollständigkeit) für eine gegebene Suchanfrage (in Anlehnung an (Baeza-Yates & Ribeiro-Neto, 1999, S. 75)).

$$\text{Recall} = \frac{\text{Anzahl der relevanten Dokumente in der Ergebnismenge}}{\text{Anzahl aller relevanten Dokumente}} \quad (5.9)$$

Abbildung 5.15 veranschaulicht die bei der Berechnung von Precision und Recall eingeschlossenen Mengen. Der Vorteil zweier verschiedener Werte liegt laut Manning, Raghavan und Schütze (2008) darin, dass einer dieser Werte je nach Anwendung des betrachteten Systems meistens wichtiger ist als der andere. Um jedoch unterschiedliche Systeme auf Basis eines einzelnen Werts miteinander vergleichen zu können, wird ein Maß benötigt, welches beide genannten Maße miteinander vereint. Das sogenannte *F-Maß* bildet das gewichtete *harmonische Mittel* von Precision und Recall.³ Standardmäßig werden Precision (P) und Recall (R) gleich gewichtet, das daraus resultierende Maß wird daher als **F₁-Maß** bezeichnet und lässt sich folgendermaßen ermitteln:

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \quad (5.10)$$

5.4.2 Vorbereitende Evaluation auf Basis von Zeitungsartikeln

Wie bereits erwähnt, wurden zwei verschiedene Evaluationsdurchläufe auf Basis von Zeitungsartikeln durchgeführt. Der erste zielte auf die Bewertung des grundsätzlichen

³Interessierte Leser seien auf Baeza-Yates und Ribeiro-Neto (1999) verwiesen.

Ansatzes, eine Themenerkennung anhand der automatischen Abbildung der in den Dialogbeiträgen genannten Terme auf die von Wikipedia bereitgestellten Kategorien zu realisieren. Mit Hilfe der zweiten Evaluation wurde schließlich die Funktionalität des gesamten, dem finalen Themenerkenners zugrunde liegenden Systems analysiert. Im Gegensatz zu dem ersten Evaluationsdurchlauf kam an dieser Stelle demnach zusätzlich der Aspekt der Satz-übergreifenden Identifikation eines Themas zum Tragen. Diese setzt das Erreichen eines zuvor festgelegten Schwellwerts an Sicherheit bezüglich der Korrektheit des identifizierten Dialogthemas voraus. Ein weiteres Ziel der zweiten Evaluation war daher die Festlegung der optimalen Höhe dieses Schwellwerts.

Evaluation des Wikipedia-Ansatzes

In einer ersten, sondierenden Evaluation wurden die innerhalb des automatischen Themenerkenners grundlegenden Prozesse der Themenidentifikation und Themenbenennung beurteilt. Dazu wurde eine Menge an natürlichsprachlichen Zeitungsartikeln der deutschen Internetzeitung *Zeit Online* auf eine vordefinierte Menge an 20 verschiedenen Wikipedia-Kategorien abgebildet. Die Kategorien entsprachen dabei den von der Zeit Online bereitgestellten Nachrichtenkategorien, in denen die einzelnen Nachrichtenartikel der Zeitung einsortiert sind. Tabelle 5.3 liefert einen Überblick über die verwendeten Kategorien.

Die gesamte Evaluation wurde in zwei Schritten durchgeführt: Zuerst wurde die korrekte Klassifikation der Themenidentifikation auf Basis von 2.000 Sätzen, welche einzelne Dialogäußerungen repräsentierten, untersucht. Als zweites wurde die grundsätzliche Korrektheit der Themenidentifikation auf Grundlage eines aus den ersten 20 Sätzen eines Nachrichtenartikels bestehenden Textabschnitts, und damit die Korrektheit hinsichtlich einer gesamten Dialogsequenz, beurteilt. Zu diesem Zweck wurden 100 Sätze von fünf unterschiedlichen Nachrichtenartikeln für jede der 20 verschiedenen, zuvor festgelegten Kategorien extrahiert. Die einzelnen Themenzugehörigkeiten der Sätze waren dabei bereits durch die von der Zeit Online bereitgestellte Zuordnung der Sätze und Artikel zu den Nachrichtenkategorien gegeben. Wie bereits erwähnt, beinhalteten diese Experimente lediglich die Prozesse zur Themenidentifikation und zur Themenbenennung. Die der laut Definition zugrunde liegende Auffassung eines Dialogthemas als gemeinsames Produkt der Dialogteilnehmer wird somit innerhalb dieser Evaluation außer Acht gelassen.

Tabelle 5.3: Evaluationsergebnisse der automatischen Themenidentifikation und Themenbenennung. Die F_1 -Werte der mittleren Spalte verdeutlichen die Korrektheit der genannten Erkennenprozesse auf Basis einzelner Artikelsätze, die Werte der rechten Spalte geben die Ergebnisse für die Erkennung auf Artikelebene wieder.

Wikipedia-Kategorie	Satz	Artikel
Beruf	0,16	0,40
Computerspiele	0,14	0,40
Essen und Trinken	0,57	1,00
Familie	0,20	0,40
Finanzierung	0,42	1,00
Geschichte	0,23	0,40
Gesundheit	0,46	1,00
Hochschule	0,29	0,80
Internationale Politik	0,19	0,40
Politik (Deutschland)	0,21	0,40
Internet	0,25	0,60
Kunst	0,13	0,20
Literatur	0,27	0,80
Medien	0,28	0,60
Mode	0,21	0,60
Musik	0,35	1,00
Umwelt und Natur	0,23	0,40
Schule	0,25	0,80
Sport	0,29	0,80
Unternehmen	0,24	0,20
\emptyset	0,27	0,61

Themenzugehörigkeit einzelner Sätze Gegenstand des erstes Experiments war die thematische Bestimmung einzelner natürlichsprachlicher Äußerungen eines Dialogs, hier repräsentiert durch die einzelnen Sätze der Nachrichtenartikel. Tabelle 5.3 zeigt die für die korrekt identifizierten Themenzugehörigkeiten ermittelten F_1 -Werte, das heißt für die Artikelsätze, die der richtigen Kategorie zugeordnet wurden. Die Ergebnisse verdeutlichen, wie unterschiedlich anspruchsvoll die Identifikation der verschiedenen, von der Internetzeitung vorgegebenen Themen ist. Insgesamt ließ sich ein F_1 -Wert von 0,27 ermitteln.

Themenzugehörigkeit ganzer Artikelabschnitte Um die Themenbestimmung anhand der ersten 20 Sätze unterschiedlicher Artikel, und somit hinsichtlich ihrer Kor-

rektheit gegenüber längeren Dialogsequenzen zu evaluieren, wurden in einem zweiten Schritt die zuvor korrekt klassifizierte Themenzugehörigkeiten der einzelnen Artikelsätze addiert. Die einzelnen Textabschnitte galten genau dann als thematisch korrekt klassifiziert, wenn die Kategorien mit den jeweils höchstem Wert mit der tatsächlichen Kategorie der Artikel übereinstimmte. Der hier erzielte F_1 -Wert (s. Tabelle 5.3) von 0,61 entspricht der Annahme, dass eine thematische Bestimmung einer größeren Anzahl an kohärenten Sätzen präziser durchzuführen ist als die Bestimmung einzelner Sätze.

Diskussion Die Ergebnisse der in dem ersten Schritt evaluierten Identifikation des Themas eines Nachrichtenartikels auf Satzebene fielen je nach Nachrichtenkategorie sehr unterschiedlich und in ihrer Gesamtheit mit einem F_1 -Wert von 0,27 deutlich schwächer als die Themenidentifikation auf Artikelebene (F_1 -Wert: 0,61) aus. Fraglich ist, ob dieses Resultat lediglich mit der gegebenen Struktur des Wikipedia-Kategoriengraphen zusammenhängt. Eine andere Erklärung wäre, dass einige Nachrichtenkategorien sehr eng miteinander in Beziehung stehen und ihre thematische Abgrenzung dadurch erschwert wird. Dies erklärt auch die Beobachtung, dass die falsch ermittelten Kategorien für die Sätze eines Artikels oftmals eine bestimmte thematische Richtung aufzeigen. Für die Sätze der eigentlich in der Kategorie „Unternehmen“ eingeordneten Artikel wurde im Durchschnitt beispielsweise häufiger das Thema „Finanzierung“ als das eigentliche Thema ermittelt, da ein Großteil der Artikel die mit der Finanzkrise einhergehenden Schwierigkeiten einzelner Unternehmen behandelt. Auch fassen bestimmte Nachrichtenkategorien unterschiedliche Themen zusammen. So beinhalten die Artikel der Kategorie „Familie“ häufig Themen wie „Gesundheit“ oder „Schule“.

Die Begründung für die erfolgreichere Bestimmung des Themas ganzer Artikelabschnitte liegt in der geringen Bedeutung, die „thematischen Ausreißern“ wie beispielsweise Nebenbemerkungen durch diese Art der Verarbeitung zukommt. Dies entspricht der Theorie nach Schank (1977) und der dieser Arbeit zugrunde liegenden Definition, wonach einzelne Sätze an sich keine eigenen Themen besitzen. Andere, Wikipedia-basierte Ansätze, zum Beispiel von Waltinger und Mehler (2009) oder Gabrilovich und Markovitch (2007), weisen diesbezüglich zwar bessere Resultate auf, lassen sich jedoch nur schwer mit dem hier verfolgten Ansatz vergleichen, da sie sich auf die Klas-

sifikation von Dokumenten als Ganzes konzentrieren. Der Unterschied liegt dabei zum einen in der Zuordnung eines einzelnen Themas für das gesamte Dokument, bei der die zusätzliche Erkennung von thematischen Wechseln innerhalb des zu klassifizierenden Textes wegfällt. Darüber hinaus laufen die bisher bekannten Erkennungsprozesse auf dem vollständigen, dem System in seiner Gesamtheit zugänglichen Dokument, welches, im Gegensatz zu der echtzeitfähigen Themenerkennung in laufenden Dialogen, keine online-fähige Verarbeitung der einzelnen Sätze erfordert. Zusammengefasst lässt sich also vermuten, dass die Identifikation eines Dialogthemas über mehrere Äußerungen hinweg erfolgversprechender ist, als die Ermittlung eines Satzthemas für jede einzelne Äußerung. Diese Hypothese wurde anhand der nachfolgend dargestellten Evaluation überprüft.

Evaluation des globalen Themenerkenners

Anhand der ersten Evaluation konnte für den generellen Ansatz, eine online-fähige Themenerkennung auf Grundlage der Wikipedia umzusetzen, eine vielversprechende Korrektheit mit einem F_1 -Wert von 0,61 für Textabschnitte ermittelt werden. Neben der Online- und Echtzeitfähigkeit des Systems bestand eine weitere Herausforderung in der Entwicklung eines Themenerkenners, dessen Erkennungsprozesse Satzübergreifend (global) arbeiten und so den Aspekt des interaktiven Aufbaus eines Dialogthemas realisieren. Der aus diesen Überlegungen hervorgegangene Themenerkenners und dessen einzelne, in Abschnitt 5.3 ausführlich vorgestellten Prozessschritte wurden anschließend in einer zweiten, ebenfalls auf dem zuvor angelegten Zeitungskorpus durchgeführten Evaluation getestet und hinsichtlich der angenommenen Verbesserung gegenüber der ersten Evaluation geprüft. Das in Echtzeit, für die einzelnen Sätze der Nachrichten ermittelte Dialogthema wird dabei durch diejenige Wikipedia-Kategorie repräsentiert, welche nach der Durchführung des Themenerkenners-Algorithmus (s. Abschnitt 5.3.3) den höchsten Wert besitzt. Die Forderung nach Überschreitung eines bestimmten Schwellwerts wurde hier demnach vernachlässigt. Tabelle 5.4 liefert einen Überblick über die Ergebnisse für die einzelnen Nachrichtenkategorien, wobei die erzielten F_1 -Werte dieses Evaluationsschrittes unter „Best“ aufgeführt sind.

Das Gesamtergebnis dieser zweiten Evaluation mit einem durchschnittlichen F_1 -Wert von 0,32 bestätigt die zuvor aufgestellte Vermutung. Demnach liefert ein Satzübergreifender Ansatz für die online-fähige Erkennung bessere Ergebnisse als die von-

Tabelle 5.4: Darstellung der für den globalen Themenerkennung im Rahmen der Evaluation erzielten F_1 -Werte.

Wikipedia Kategorie	<i>Best</i>	$i = 0, 2$	$i = 0, 3$	$i = 0, 35$	$i = 0, 4$
Beruf	0,04	0,03	0,04	0,04	0,04
Computerspiele	0,23	0,25	0,25	0,21	0,21
Essen und Trinken	0,62	0,36	0,55	0,66	0,62
Familie	0,06	0,05	0,04	0,05	0,05
Finanzierung	0,51	0,46	0,53	0,53	0,50
Geschichte	0,49	0,33	0,36	0,29	0,23
Gesundheit	0,68	0,68	0,71	0,70	0,71
Hochschule	0,18	0,15	0,16	0,16	0,15
Internationale Politik	0,22	0,15	0,20	0,22	0,24
Politik (Deutschland)	0,23	0,17	0,24	0,25	0,22
Internet	0,20	0,22	0,16	0,16	0,14
Kunst	0,15	0,0	0,07	0,10	0,10
Literatur	0,42	0,24	0,36	0,29	0,22
Medien	0,50	0,06	0,37	0,31	0,20
Mode	0,16	0,10	0,12	0,12	0,13
Musik	0,49	0,03	0,42	0,34	0,20
Umwelt und Natur	0,42	0,36	0,43	0,43	0,42
Schule	0,11	0,10	0,11	0,11	0,10
Sport	0,54	0,15	0,50	0,57	0,51
Unternehmen	0,23	0,23	0,18	0,18	0,18
\emptyset	0,32	0,21	0,29	0,29	0,26

einander unabhängige Verarbeitung einzelner Sätze, für die ein durchschnittlicher F_1 -Wert von 0,27 ermittelt wurde. Darüber hinaus konnte die Fehleranfälligkeit des Erkenners, welche bei der ersten Evaluation bei 0,73% lag und daher mit der Anzahl der nicht korrekt erkannten Themen übereinstimmt, minimiert werden. Genauer gesagt wurde in 152 Fällen (7,6%) der 1.385 Sätze, die nicht korrekt klassifiziert wurden, kein Thema zurückgeliefert. Im Hinblick auf die Zielanwendung des hier evaluierten Themenerkenners in der Mensch-Maschine-Interaktion, ist die weitere Reduzierung dieser Fehleranfälligkeit von hoher Wichtigkeit, um die Glaubwürdigkeit des mit dem Themenerkennung ausgestatteten Agenten Max zu erhöhen. Diese wird dadurch erreicht, indem eher *kein* Thema als ein falsches erkannt wird. Aus diesem Grund wurde ein **Schwelwert** i eingeführt, welcher die Anzahl der nicht korrekt identifizierten Dialogthemen minimiert. Eine Koordinate innerhalb des Dialogthemen-Vektors muss demnach einen

Tabelle 5.5: Fehleranfälligkeit des automatischen Themenerkenners je nach Höhe des Schwellwerts.

Schwellwert	Fehleranfälligkeit
Best	61,7%
$i = 0,2$	24,0%
$i = 0,3$	36,2%
$i = 0,35$	33,1%
$i = 0,4$	23,1%

bestimmten Mindestwert erreichen, damit die Kategorie, welche durch die Koordinate repräsentiert wird, als Dialogthema identifiziert werden kann (s. Abschnitt 5.3.2).

Schwellwert-Bestimmung Mit dem Ziel, diesen Schwellwert für die zukünftige Anwendung des Themenerkenners festzulegen, wurden weitere, wie oben beschriebene Evaluationsdurchläufe vorgenommen, hier jedoch mit unterschiedlich voreingestellten Schwellwerten. Die einzelnen Schwellwerte lagen dabei, entsprechend der möglichen Kategorienwerte, zwischen 0 und 1. Abbildung 5.4 präsentiert die besten der aus der Evaluation hervorgegangenen F_1 -Werte für die unterschiedlichen Schwellwerte. Zusätzlich zu diesen Werten wurde darüber hinaus die jeweilige Fehleranfälligkeit des Systems beurteilt. Dazu wurden sowohl die richtig klassifizierten Äußerungen als auch die, für die kein Thema identifiziert werden konnte, als richtige Antworten definiert. Dabei wurde auch dann „kein“ Thema ermittelt, wenn mehr als nur ein aktuelles Dialogthema identifiziert wurde und somit ein mehrdeutiges Ergebnis vorlag. Tabelle 5.5 listet die resultierenden Prozentwerte für die Fehleranfälligkeit je nach Schwellwert-Bestimmung auf. Anhand dieser Werte ist ersichtlich, dass der Themenerkennner ohne Schwellwert-Bestimmung (61,7% für „Best“) bedeutend fehleranfälliger ist und die Festlegung eines Schwellwerts für die Identifikation eines Themas daher unerlässlich ist.

Zuvor wurde der höchste F_1 -Wert (0,29) für einen Schwellwert von $i = 0,3$ und $i = 0,35$ erreicht (s. Tabelle 5.4). Demgegenüber steht jedoch ein erkennbarer Unterschied in der Fehleranfälligkeit des Systems hinsichtlich dieser beiden Schwellwerte. Mit $i = 0,3$ wurden in 36,2% der Prozessdurchläufe falsche Ergebnisse zurückgeliefert, mit einem Schwellwert von $i = 0,35$ hingegen lag die Fehleranfälligkeit bei 33,1%. Aus diesem Grund wurde der Schwellwert für die Themenerkennung auf $i = 0,35$ festgelegt.

5.4.3 Evaluation auf Basis des CUBE-G-Korpus

Die Ergebnisse der bis hierher beschriebenen Evaluationsdurchgänge erlaubten einen ersten Einblick in die Funktionalität des Themenerkenners. Immerhin gelang mit dem Zeitungskorpus die Festlegung des Schwellwerts für die Themenbestimmung. Mit dem Erhalt des *CUBE-G*-Korpus wurde schließlich eine der eigentlichen Zielanwendung des automatischen Themenerkenners vergleichbare Evaluationsgrundlage erworben.

Das CUBE-G-Korpus

Das von Endrass, Rehm und André (2011) erstellte **CUBE-G**-Korpus besteht aus annotierten Aufnahmen von insgesamt 21 Kennenlerngesprächen zwischen zwei deutschsprachigen Dialogpartnern mit einer durchschnittlichen Länge von fünf Minuten. Die Gespräche fanden im Rahmen einer vermeintlichen Studie in der Universität Augsburg statt, für die sich die bis dahin unbekanntenen Probanden im Vorfeld kennenlernen sollten. Die dabei entstandenen Aufnahmen lieferten die Grundlage für die Erforschung kultureller Unterschiede in Smalltalk-Situationen, auf dessen Basis die Entwicklung Kultur-spezifischer konversationaler Agenten praktiziert wird. Im Rahmen der Beobachtung typischer Verhaltensweisen wurden die Dialoge unter anderem hinsichtlich der Anzahl besprochener Themen und vollzogener Themenwechseln untersucht. Die einzelnen Äußerungen der Dialogteilnehmer wurden zu diesem Zweck mit insgesamt acht unterschiedlichen Themen annotiert.

Online-Studie

Für die Evaluation des Themenerkenners hinsichtlich seines Einsatzes in natürlichen Dialogen war ursprünglich ein Abgleich der automatisch zugewiesenen Dialogthemen mit den manuellen Annotationen geplant. Die ersten Ergebnisse verdeutlichten jedoch auch hier die Problematik der thematischen Abgrenzung. So wurde beispielsweise einer Dialogsequenz über die Berufsperspektive eines Informatik-Studenten manuell das Thema „Studium“ zugewiesen, wohingegen der automatische Themenerkenners für dieselbe Sequenz das Thema „Beruf“ ermittelte. Aus subjektiver Sicht sind beide Zuordnungen richtig, sodass eine Bewertung der Korrektheit der Themenerkennung durch Abgleich der automatischen und manuellen Themenzuweisungen als unzureichend erachtet wurde. Stattdessen wurde eine Online-Studie durchgeführt, bei der die

Tabelle 5.6: Mögliche Themen der CUBE-G-Dialoge und ihre Kategorienrepräsentationen.

Dialogthema	Wikipedia-Kategorie
Studium	Studium
Alter	Alter
Freizeit	Freizeit, Hobby
Tourismus	Tourismus, Urlaub und Erholung
Wohnort	Orte in Deutschland
Beruf	Beruf
Gesundheit	Gesundheit
Musikinstrument	Musikinstrument
Informatik	Informatik

Studienteilnehmer die Ergebnisse der automatischen Themenzuweisung beurteilten.

Vorverarbeitung In einem ersten Schritt galt es eine Menge an vordefinierten Wikipedia-Kategorien festzulegen, welche die möglichen Themen für die betrachtete Dialogsituation repräsentierten. Die Grundlage hierzu lieferten die Themen, mit denen die Dialoge innerhalb des CUBE-G-Korpus manuell annotiert wurden, jedoch wurde zusätzlich das Thema „Informatik“ hinzugefügt, da sich aus subjektiver Sicht weite Teile der einzelnen Dialoge mit diesem Thema befassten. Tabelle 5.6 liefert einen Überblick über die resultierende Themenmenge und die entsprechenden Wikipedia-Kategorien. Daraus wird ersichtlich, dass die Themen „Freizeit“ und „Tourismus“ durch jeweils zwei Kategorien repräsentiert werden. Der Grund hierfür liegt in der vollständigeren Abbildung des jeweiligen Themas. Im Anschluss an die Festlegung der vordefinierten Themenmenge wurden ein entsprechender *Kategorienindex* (s. Abschnitt 5.1.2) sowie ein aktueller *Artikelindex* aufgebaut.

In einem zweiten Vorverarbeitungsschritt wurde nun die automatische Themenerkennung durchgeführt. Dafür wurden zunächst die einzelnen Dialogbeiträge den Bedingungen der Tastatur-basierten Eingabe angepasst. Wiederholungen und unvollständige Sätze wurden dabei zusammengefasst und Bemerkungen wie „Äh“ etc. entfernt. Anschließend wurden die einzelnen Dialoge wie in Abschnitt 5.3 dargestellt verarbeitet und jede einzelne Äußerung mit dem jeweils aktuellen Dialogthema annotiert. Ließ sich aufgrund zu geringer Relevanzwerte oder aufgrund von Mehrdeutigkeiten kein Thema bestimmen, wurde die entsprechende Äußerung mit einem Bindestrich annotiert.

Durchführung Für die Durchführung der Studie wurde eine Online-Umfrage bestehend aus vier Einzelumfragen erstellt. Jede Einzelumfrage umfasste fünf der 21 verarbeiteten CUBE-G-Dialoge, sodass der letzte Dialog des Korpus von der Umfrage ausgeschlossen und somit nicht bewertet wurde. Dem jeweiligen Studienteilnehmer wurden die Dialoge wie in Abbildung 5.16 dargestellt präsentiert. Das heißt, jede einzelne Äußerung sowie das dieser Äußerung zugeordnete Dialogthema wurden in einer separaten Zeile abgebildet. Die Aufgabe der Studienteilnehmer bestand nun in der Beurteilung, ob das jeweils angegebene Thema passend für die jeweilige Dialogsituation ist. Die Beurteilung wurde dabei durch Anklicken von „Ja“ oder „Nein“ rechts neben der jeweiligen Äußerung durchgeführt. Äußerungen ohne Themenangabe sollten dahingehend bewertet werden, ob für die jeweilige Dialogsituation tatsächlich kein Thema vorliegt.

Jeder Studienteilnehmer bearbeitete eine Einzelumfrage und bewertete somit die Ergebnisse der automatischen Themenerkennung für fünf Dialoge. Insgesamt nahmen 11 Frauen und 4 Männer (insgesamt also 15 Personen) zwischen 28 und 64 Jahren an der Studie teil.

Ergebnisse Die Auswertung der Studie wurde auf Basis unterschiedlicher Berechnungen vorgenommen. Zur Vereinfachung ihrer Darstellung sollen zunächst die betrachteten Ergebnismengen definiert werden:

- **D** definiert die Menge aller Dialogzeilen
- **T** definiert die Menge aller Dialogzeilen, denen mittels der automatischen Themenerkennung ein Dialogthema zugewiesen wurde
- **nT** definiert die Menge aller Dialogzeilen, denen mittels der automatischen Themenerkennung *kein* Dialogthema zugewiesen wurde
- **Y** definiert die Menge aller Dialogzeilen, die von den Studienteilnehmern mit „Ja“ bewertet wurden
- **N** definiert die Menge aller Dialogzeilen, die von den Studienteilnehmern mit „Nein“ bewertet wurden

Es gilt also $D = T + nT$ und $D = Y + N$. Ferner ist zu beachten, dass jede Dialogzeile so oft in den jeweiligen Mengen enthalten ist, wie sie von den Studienteilnehmern

Dialog 1

Im Folgenden sehen Sie ein Kennenlerngespräch zwischen zwei Personen. Jede Äußerung dieses Dialogs ist mit einer Themenzuweisung versehen.
Bitte beurteilen Sie, ob das jeweils angegebene Thema passend für die jeweilige Dialogsituation ist. Bei Äußerungen ohne Themenangabe entscheiden Sie bitte, ob diese Angabe passt und aktuell tatsächlich kein Thema vorliegt.

		Ja	Nein
Ich heiße Mark.	--	<input type="radio"/>	<input type="radio"/>
Ich heiße Stefan.	--	<input type="radio"/>	<input type="radio"/>
Okay.	--	<input type="radio"/>	<input type="radio"/>
Was machst du so?	--	<input type="radio"/>	<input type="radio"/>
Ich studiere Theatergeschichte in München.	--	<input type="radio"/>	<input type="radio"/>
Ich studiere auch in München, ich studiere Lehramt Bio und Chemie. Bin jetzt gerade mitten im Staatsexamen eigentlich.	Studium	<input type="radio"/>	<input type="radio"/>
Aber du bist eigentlich Augsburgener?	Studium	<input type="radio"/>	<input type="radio"/>
Ich komme aus Augsburg, aber meine Freundin wohnt hier. Die arbeitet auch hier, deshalb wollte ich auch nicht nach München ziehen. Bist du dann auch gependelt die ganze Zeit?	Wohnort	<input type="radio"/>	<input type="radio"/>
Nein, ich habe schon eine schöne Wohnung in München.	Wohnort	<input type="radio"/>	<input type="radio"/>
Ach so. Aber am Wochenende halt oder so?	Wohnort	<input type="radio"/>	<input type="radio"/>
Ja genau, oder auch mal unter der Woche, wenn die Sehnsucht einen treibt.	Wohnort	<input type="radio"/>	<input type="radio"/>
Das geht ja, es sind ja keine Entfernungen, man fährt ja nur eine halbe Stunde mehr oder weniger. Das ist ja eigentlich ganz praktisch.	Wohnort	<input type="radio"/>	<input type="radio"/>
Aber so auf Dauer pendeln, das war schon immer anstrengend, jeden Tag dann. Manchmal hätte ich mir gewünscht, eine Wohnung drüben zu haben, aber im Endeffekt war es dann doch immer schön, wenn man abends hier ist.	Wohnort	<input type="radio"/>	<input type="radio"/>
Fährst du mit dem Zug oder mit dem Auto?	--	<input type="radio"/>	<input type="radio"/>
Mit dem Zug.	Freizeit	<input type="radio"/>	<input type="radio"/>
		Ja	Nein
Man kann die Zugfahrt eigentlich immer ganz gut nutzen für irgendwas lesen oder so.	Freizeit	<input type="radio"/>	<input type="radio"/>

Abbildung 5.16: Auszug aus der Online-Studie.

bewertet wurden. Beispiel: Wenn vier Nutzer den Fragebogen über Dialog 1 ausgefüllt haben, sind alle Dialogzeilen von Dialog 1 vier Mal in den entsprechenden Mengen vorhanden.

Tabelle 5.7: Ergebnisse der Online-Studie.

Akkuratheit:	43,7%
Fehleranfälligkeit:	28,7%
F_1 :	0,40

Berechnet wurde zum einen die allgemeine Akkuratheit ($|Y|/|D|$) und die Fehleranfälligkeit ($|T \cap N|/|D|$) des Themenerkenners. Die Akkuratheit beinhaltet dabei auch die Dialogzeilen ohne Themenangabe, sodass diese wie ein zusätzliches Thema bewertet wurden. Für die Berechnung von F_1 hingegen wurden nur die tatsächlichen Themenzuweisungen betrachtet, da diese Betrachtungsweise der Auswertung des Zeitungskorpus entspricht und somit eine bessere Vergleichsbasis zugrunde liegt. Dementsprechend wurden für Precision und Recall (s. Gleichungen 5.8 und 5.9) die Dialogzeilen als *relevante Dokumente* definiert, für die eine Themenzuweisung hätte erfolgen sollen, also $T + (nT \cap N)$. *Alle Dokumente in der Ergebnismenge* entsprach dann allen Dialogzeilen D . Die *relevanten Dokumente in der Ergebnismenge* waren dann die Dialogzeilen, die ein Thema zugewiesen bekommen hatten und dieses zusätzlich von den Studienteilnehmern als richtig bewertet wurde, also $T \cap Y$. Insgesamt ergaben sich daraus die in Tabelle 5.7 zusammengetragenen Werte.

Diskussion

Die Ergebnisse der Online-Studie belegen die erwartete Verbesserung der automatischen Themenerkennung für dessen Einsatz in natürlichen Dialogen. So liegt die im Rahmen der Online-Studie ermittelte Akkuratheit des Themenerkenners mit insgesamt 43,7% deutlich über der für die Zeitungsevaluation ermittelte Akkuratheit von insgesamt 23,3% (die Akkuratheit entsprach in diesem Fall dem zuvor ermittelten Recall-Wert). Gründe hierfür liegen vermutlich in der Tatsache, dass Dialoge durch häufigere Themenwechsel gekennzeichnet sind, welche sich durch den Themenerkennner ermitteln lassen und zu der Identifikation unterschiedlicher Themen für einen Dialog führen. Darüber hinaus verdeutlicht die deutliche Steigerung der Akkuratheit zudem den Aspekt der *Subjektivität* bei der Zuordnung und Bezeichnung von Dialogthemen. So erzielte eine subjektive Bewertung der Ergebnisse eine höhere Akkuratheit als ein Abgleich der Ergebnisse mit bereits annotierten Themen.

Besonders erfreulich ist die niedrigere Fehleranfälligkeit der Themenerkennung in

natürlichen Dialogen, im Hinblick auf die damit einhergehende Steigerung der Glaubwürdigkeit des mit dem Themenerkennner ausgestatteten Agenten Max. Diese lässt sich auch anhand des F_1 -Werts nachweisen, welcher mit 0,40 eine Erhöhung von 0,11 gegenüber dem zuvor für die Evaluation auf Basis des Zeitungskorpus ermittelten Wert aufweist. Die relativ geringere Steigerung dieses Werts im Gegensatz zu der allgemeinen Verbesserung der Akkuratheit lässt sich dabei durch den Ausschluss der Dialogzeilen ohne Themenangabe erklären. Die Mitbetrachtung dieser Zuweisungen ergibt einen F_1 -Wert, der mit der allgemeinen Akkuratheit übereinstimmt, also 0,44.

Insgesamt lässt sich anhand der hier dargestellten Ergebnisse eine erhöhte Adäquatheit des Themenerkenners hinsichtlich seiner Anwendung in natürlichen Dialogsituationen nachweisen. Gleichzeitig verdeutlichen die Ergebnisse allerdings auch die Notwendigkeit weiterer Arbeitsschritte zur Verbesserung der Themenerkennung. Dabei sollte in weiterführenden Arbeiten insbesondere die Auflösung mehrdeutiger Ergebnisse in Form von mehreren Dialogthemen umgesetzt werden. Darüber hinaus sollten in Zukunft weitere Evaluationen auf Basis des CUBE-G-Korpus angestrebt werden. Ein Ansatz dabei wäre beispielsweise die Definition neuer Themenmengen, um den Unterschied in der Betrachtung speziellerer und allgemeinerer Themen zu untersuchen. Ein weiterer Aspekt für zukünftige Evaluationsschritte wäre die Bewertung automatisch erkannter Themenwechsel.

5.5 Zusammenfassung

In diesem Kapitel wurde die technische Umsetzung des in Kapitel 4 konzipierten Themenmodells beschrieben. Zu Beginn wurde die Indexierung der in der Wikipedia enthaltenen, für die automatische Themenerkennung erforderlichen Informationen auf Basis von Lucene dargestellt. Diese lieferte die Basis für den Aufbau eines künstlichen mentalen Lexikons, welches einen schnellen maschinellen Zugriff auf Wissen über Konzepte, repräsentiert durch Wikipedia-Artikel, und ihre thematischen Zusammenhänge in Form von Wikipedia-Kategorien erlaubt.

Die nachfolgenden Abschnitte beschäftigten sich ausführlich mit der Präsentation derjenigen Komponente des Themenmodells, welche die automatische Erkennung von Themen in laufenden Dialogen durchführt, dem Themenerkennner. Dieser war so umzusetzen, dass simultan zu dem eigentlichen Prozess der Themenerkennung weitere

Aufgaben zur Identifikation von Informationen über den thematischen Ablauf des betrachteten Dialogs durchgeführt werden. Gleichzeitig musste der Themenerkennung den Forderungen nach Generalität und Echtzeitfähigkeit nachkommen. Als Ergebnis wurde ein Algorithmus entwickelt, deren einzelne Prozesse die Ableitung der Themeninformationen und die Bestimmung des aktuellen Dialogthemas durch eine dynamische, online-fähige und globale Themenerkennung ermöglichen.

Zum Schluss des Kapitels wurden die unterschiedlichen Evaluationsdurchgänge vorgestellt, die Aufschluss über die grundsätzliche Lauffähigkeit des Wikipedia-basierten Ansatzes, die Bestimmung des Schwellwerts zur Identifikation einer Kategorie als Dialogthema sowie Aufschluss über den Einsatz des Themenerkenners in natürlichen Dialogsituationen gaben. Das nächste Kapitel liefert eine Beschreibung, wie die Integration dieser Komponente in das Konversationsverhalten des Agenten Max die Emulation menschenähnlichen Themenbewusstseins in Dialogen zwischen Mensch und Agent realisiert.

6 Integration des Themenmodells

Im vorherigen Kapitel wurde ein Modell für die automatische Erkennung von Themeninformationen in natürlichen Dialogen erarbeitet. Um darauf aufbauend menschenähnliches Themenbewusstsein in einem konversationalen System emulieren zu können, bedarf es zweier Integrationsschritte: Erstens muss die Einbettung des Themenerkenners in die bestehende Systemarchitektur des jeweiligen Zielsystems umgesetzt werden. Zweitens müssen die während des Dialogs mit Hilfe des Themenerkenners erworbenen Themeninformationen in das konversationale Verhalten des Agenten einbezogen werden. Das vorliegende Kapitel stellt die Realisierung dieser, in Abschnitt 4.4 bereits konzipierten, Integrationsmaßnahmen innerhalb des konversationalen Agenten *Max* vor. Wie in Abbildung 6.1 veranschaulicht, wird der automatische Themenerkennner des Themenmodells zu diesem Zweck derart an die Systemarchitektur des Agenten angebunden, dass die aus dem Erkennungsprozess resultierenden **Themeninformationen** in die das konversationale Verhalten des Agenten realisierende **Deliberative Komponente** einfließen und auf diese Weise die Generierung der **Dialogbeiträge** für den Agenten beeinflussen.

Dieses Kapitel gliedert sich folgt: Es beginnt mit einer Vorstellung des Zielsystems, deren Fokus auf der Beschreibung der bestehenden Dialogfähigkeiten des Agenten liegt. Im Anschluss wird in Abschnitt 6.2 die Einbettung des Themenerkenners in die durch ein Multiagentensystem realisierte Gesamtarchitektur des konversationalen Agenten vorgestellt. Anschließend, in Abschnitt 6.3, erfolgt die Darstellung der Anpassung des Dialogwissens von Max, welche eine Integration der gewonnenen Themeninformationen in das konversationale Verhalten des Agenten und somit die Ausführung thematisch bedingter Äußerungen ermöglicht. Mit dem Ziel, das aus den beschriebenen Integrationsschritten realisierte künstliche Themenbewusstsein von Max aufzuzeigen, wird abschließend, in Abschnitt 6.4, ein konkretes Anwendungsbeispiel beschrieben.

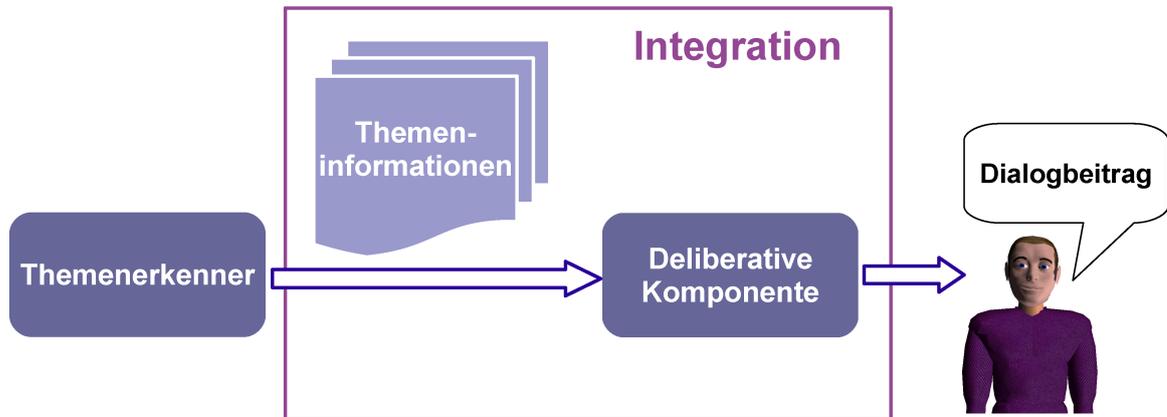


Abbildung 6.1: Darstellung der Einbettung des automatischen Themenerkennners in die Architektur des konversationalen Agenten *Max*, auf dessen Basis die Integration aktueller *Themeninformationen* in die *Deliberative Komponente* des Agenten gelingt. Ziel ist die Generierung von *Dialogbeiträgen* des Agenten, welche die für einen Dialog zwischen Menschen typische Fähigkeit, sich des thematischen Inhalts und der unterschiedlichen, den thematischen Ablauf eines Dialogs beeinflussenden Aktionen bewusst zu sein, emulieren.

6.1 Der konversationale Agent Max

Max ist ein virtueller, konversationaler Agent. Je nach Anwendungsszenario ist es menschlichen Interaktionspartnern möglich, durch gesprochene oder textuelle Eingaben mit Max zu kommunizieren. Dabei reagiert der Agent nicht nur auf die Eingaben seines jeweiligen Gegenübers, sondern verhält sich in angemessenen Situationen *proaktiv*, das heißt er ergreift die Initiative und fordert seinen Dialogpartner beispielsweise auf, ein Gespräch mit ihm zu beginnen. Einsatz findet Max unter anderem als Rezeptionist auf dem Flur der Arbeitsgruppe Wissensbasierte Systeme¹ der Universität Bielefeld, wo er Besucher des Labors für künstliche Intelligenz und virtuelle Realität begrüßt und sich mit ihnen unterhält. In einem anderen Anwendungsszenario agiert der Agent als Museumsführer in dem Paderborner Heinz Nixdorf MuseumsForum² (s. Abbildung 6.2), wo dieser den Besuchern des Museums auf Anfrage Informationen über verschiedene Ausstellungsstücke gibt (Kopp et al., 2005).

¹<http://www.techfak.uni-bielefeld.de/ags/wbski/>

²www.hmf.de



Abbildung 6.2: Max als Museumsführer im Paderborner Heinz Nixdorf MuseumsForum.

6.1.1 Dialogfähigkeit

Wie bereits in Abschnitt 4.4.2 beschrieben, wird das Dialogsystem von Max durch eine **Deliberative Komponente** realisiert. Alle Prozesse dieser Komponente bauen auf der **Belief-Desire-Intention** (BDI)-Theorie auf. Das heißt, der Agent verfügt über Annahmen (*Beliefs*) über den aktuellen Zustand seiner Umwelt, über Ziele (*Desires*), die gewünschte Umweltzustände des Agenten beschreiben, und über Absichten (*Intentions*), die in Form von Plänen angeben, wie die jeweiligen Ziele zu erreichen sind. Realisiert wird dieser Ansatz auf Basis von *JAM* (Huber, o. J.). *JAM* ist eine Agentenarchitektur, die unter anderem aus einem Weltmodell und einer Planbibliothek besteht. Die Auswahl eines Plans hängt von der Anwendbarkeit und der Nützlichkeit, welche durch einen sogenannten *Utility*-Wert angegeben wird, des Plans ab und geschieht auf der Grundlage von **Regeln**.

Die Grundeinheit der konversationalen Interaktion zwischen Max und seinem menschlichen Gegenüber bilden *kommunikative Akte*, deren zugrunde liegende Theorie auf Poggi und Pelachaud (2000) zurückzuführen. Jeder sprachlichen Äußerung, sowohl der des menschlichen als auch der des künstlichen Dialogteilnehmers, wird im Rahmen der Interpretation mindestens ein kommunikativer Akt zugeschrieben, der jeweils eine bestimmte **konversationale Funktion** erfüllt. Die konversationale Funktion umschreibt dabei die funktionale Bedeutung des jeweiligen kommunikativen Akts, und beinhal-

tet drei voneinander unabhängige Komponenten, die die unterschiedlichen Aspekte der Funktion einer Äußerung angeben. Da der Dialogmanager intern auf Basis der konversationalen Funktion anstatt mit dem Text der betrachteten Äußerung arbeitet, genügt die Definition einer kleinen Menge an Regeln, die als Bedingung die jeweilige konversationale Funktion angeben. So können viele unterschiedliche Eingaben verarbeitet und passende, in den jeweiligen Regeln definierte Ausgaben für den Agenten definiert werden. Letzteren wird innerhalb der Regeln ebenfalls eine konversationale Funktion zugeschrieben, sodass sowohl die Äußerungen des Menschen als auch die des konversationalen Agenten funktional analysiert vorliegen.

Aufgrund der regelbasierten Eingabeninterpretation, die ein breites Spektrum an möglichen Äußerungen abdeckt, und der zusätzlichen, ebenfalls auf Grundlage der Wikipedia realisierten Frage-Antwort-Schnittstelle (s. Abschnitt 3.3.1), gelingt es dem Agentensystem immer, eine passende Antwort auf eine von dem menschlichen Gegenüber verfasste Eingabe zu generieren. Nichtsdestotrotz scheitert das System bisher daran, ein kohärentes Gespräch mit dem Menschen zu führen, da der Agent bislang nicht über ein künstliches Themenbewusstsein verfügt. Diese Unzulänglichkeit gilt es mit der Anbindung des Systems an das in Kapitel 5 erarbeitete Themenmodell zu beheben.

6.2 Integration des Themenerkenners in die Systemarchitektur

Das dem Agenten Max zugrunde liegende System basiert auf einem sogenannten **Multiaagentensystem**, welches sich aus mehreren miteinander interagierenden Elementen, den sogenannten *Software-Agenten*, zusammensetzt. Laut Wooldridge (2009) zeichnen sich Software-Agenten dadurch aus, dass sie zum einen dazu in der Lage sind, autonom zu handeln, und zum anderen mit weiteren Software-Agenten mittels Kommunikationsschnittstellen interagieren können. Das nach außen für den menschlichen Dialogpartner sichtbare Verhalten von Max entwickelt sich somit aus dem Verhalten und der Kooperation einzelner Agenten. Auch die das verbale Verhalten des Agenten Max realisierende Deliberative Komponente wird durch so einen intelligenten Agenten, dem sogenannten *Dialogagenten*, umgesetzt. Demnach erfolgt die Integration des Themenerkenners durch Implementierung eines entsprechenden **Themenagenten**, der wie

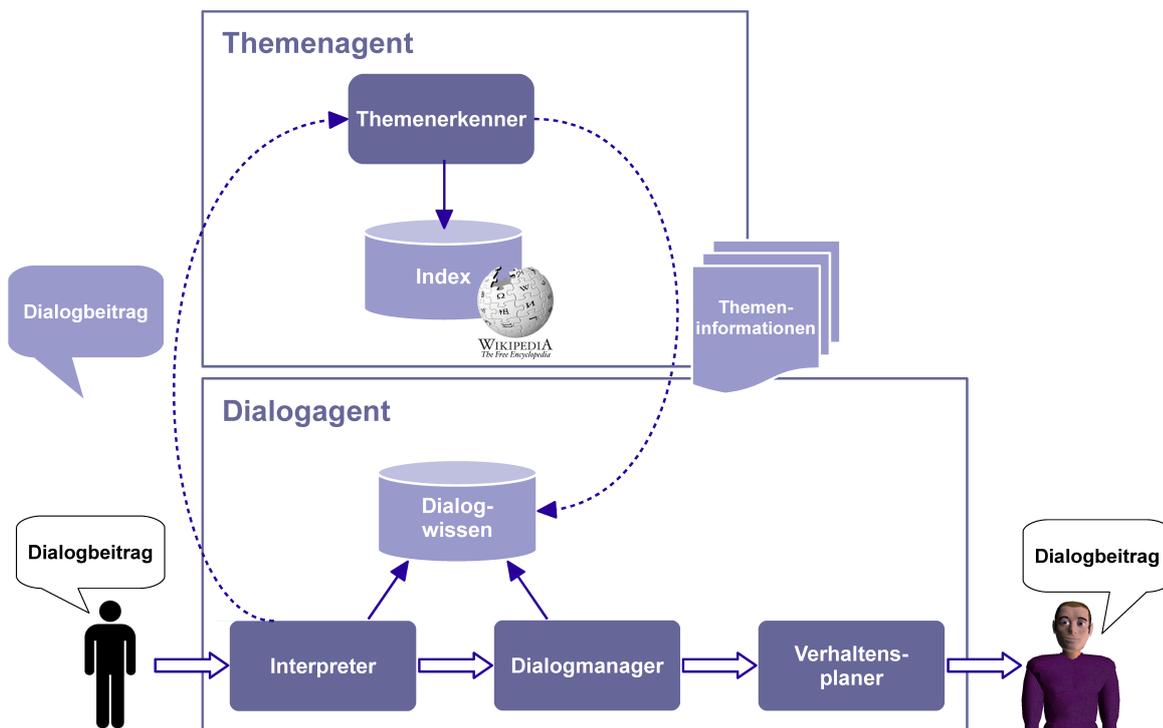


Abbildung 6.3: Für eine Integration des Themenerkenners in die Gesamtarchitektur des konversationalen Agenten Max ist die Implementierung eines *Themenagenten* erforderlich, der mit dem *Dialogagenten* kommunizieren und auf diese Weise jederzeit die aktuellen Themeninformationen an diesen weiterleiten kann.

in Abbildung 6.3 dargestellt an den Dialogagenten angebunden wird. Die bestehende **Echtzeitfähigkeit** des Systems bleibt auf diese Weise erhalten.

Jede neue Äußerung eines Dialogs (in der Abbildung aktuell die Äußerung des menschlichen Dialogpartners) erreicht zunächst den Dialogagenten. Dieser ermittelt deren konversationale Funktion und leitet das Ergebnis mitsamt der Äußerung an den Themenagenten weiter. Dieser entscheidet auf Grundlage der konversationalen Funktion über die thematische Relevanz der Äußerung und somit über deren Weiterverarbeitung. Als Begrüßung oder Verabschiedung spezifizierte Äußerungen durchlaufen die Prozesse der Themenerkennung beispielsweise nicht, da sie keinen thematischen Beitrag zu dem laufenden Dialog leisten. Wird eine Äußerung als thematisch relevant identifiziert, führt dies zu der Ausführung der Themenerkennung und zu einer entsprechenden Aktualisierung der Themeninformationen. Diese werden im Anschluss an den Dialogagenten geschickt, welcher die Informationen in das Dialogwissen einträgt. Von

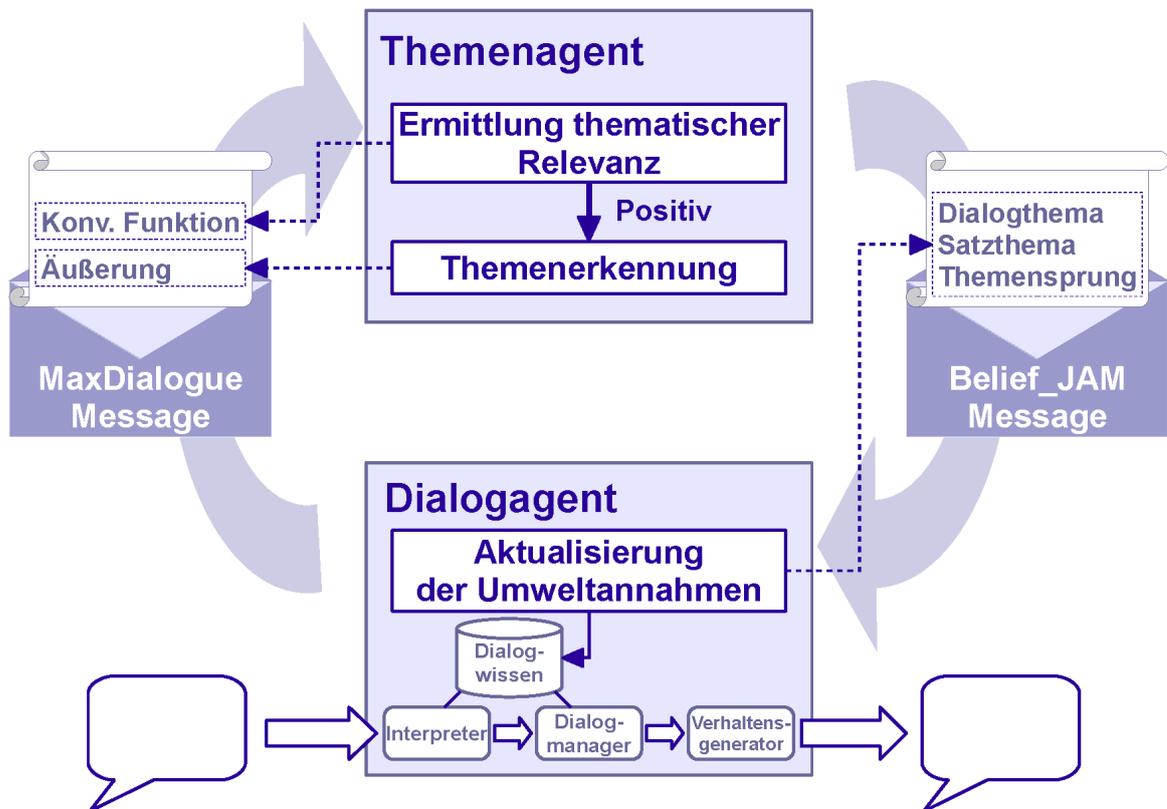


Abbildung 6.4: Kommunikation zwischen Themen- und Dialogagent zum Austausch von Äußerungen und Themeninformationen.

dort aus nehmen diese dann Einfluss auf die von dem Dialogmanager durchgeführte Auswahl einer geeigneten Verhaltensweise für Max.

Die Kommunikation zwischen Themen- und Dialogagent basiert auf dem in Abbildung 6.4 skizzierten Austausch von **Nachrichten** und liefert die Grundlage für die Integration der Themeninformationen in das konversationale Verhalten von Max. Den Agenten stehen dabei unterschiedliche Typen einer Nachricht zur Verfügung, um verschiedene Inhalte verschicken und bei dem Empfänger bestimmte Prozesse auslösen zu können. So leitet der Dialogagent die konversationale Funktion und den Inhalt der aktuellen Äußerung beispielsweise anhand einer `MaxDialogueMessage` an den Themenagenten weiter, der anhand des Typs dieser Nachricht Rückschlüsse auf die darin enthaltene Information ziehen kann. Empfängt der Themenagent also eine Nachricht vom Typ `MaxDialogueMessage`, so ermittelt dieser auf Basis der konversationalen Funktion die thematische Relevanz der Äußerung und führt erst mit Nachweis dieser

Relevanz den Prozess der Themenerkennung auf Basis dieser Äußerung durch. Die resultierenden Themeninformationen werden anschließend anhand einer Nachricht des Typs `Belief_JAMMessage` an den Dialogagenten zurückgeschickt, der daraufhin die in dem Dialogwissen enthaltenen Annahmen (*Beliefs*) über die Umwelt des Agenten mittels *JAM*-Plänen aktualisiert.

6.3 Integration der Themeninformationen in das Konversationsverhalten

Nach der physikalischen Integration des Themenerkenners in die bestehende Architektur von Max kann die Integration der Themeninformationen in das Konversationsverhalten des Agenten vorgenommen werden. Dieser Abschnitt erläutert, wie das System das anhand des Themenerkenners erlangte Wissen über die thematischen Abläufe des laufenden Dialogs in das konversationale Verhalten des Agenten einbindet und auf diese Weise menschenähnliches Themenbewusstsein emuliert. Grundlage für die genaue Umsetzung themenbewusster Verhaltensweisen liefern die in Abschnitt 4.4.3 zusammengetragenen Konzeptentscheidungen. Die technische Voraussetzung dazu liefert die in Abbildung 6.4 dargestellte Kommunikation zwischen dem Dialog- und dem Themenagenten, welche zu jedem Zeitpunkt die Aktualität der Themeninformationen gewährleistet.

6.3.1 Emulierte Fähigkeit zur Themenbenennung

Menschen sind dazu in der Lage, die in einer Konversation angesprochenen Themen zu benennen. Diese Fähigkeit liefert die Basis, das eigene Bewusstsein hinsichtlich Dialogthemen und ihren Zusammenhängen nach außen hin zu kommunizieren. Eine Emulation dieser Fähigkeit ist daher grundlegende Voraussetzung für alle weiteren Implementierungen. Wie in Abschnitt 5.3.4 geschildert, wird durch Ausführung der fünften Aufgabe an den Themenerkennner eine Themenbezeichnung festgelegt. Diese repräsentiert dann die Information über das aktuelle Dialogthema und wird mit der `MaxDialogueMessage` an den Dialogagenten geschickt (s. Abbildung 6.4). Dieser trägt die neue Information durch Ausführung des *JAM*-Plans *set-topic* (s. Algorithmus 3, Zeile 1 – 6) als *Belief*, also als aktuelle Annahme über die Umwelt, in das Dialogwissen

des Agenten ein. Bei Bedarf greift das System dann mit Hilfe des Plans *get-topic* (s. Algorithmus 3, Zeile 7 – 15) auf diese Information zu, sodass Max anhand dieser das aktuelle Dialogthema in seinen Äußerungen benennen kann. Werden gleich zwei Dialogthemen identifiziert, bleibt die entsprechende Annahme über die Umwelt leer, um Mehrdeutigkeiten ausschließen zu können.

Algorithmus 3 JAM-Pläne zum Eintragen (*set-topic*) und Abfragen (*get-topic*) des aktuellen Dialogthemas in das bzw. aus dem Dialogwissen des Agenten. Die Variable *\$topic* repräsentiert dabei den entsprechenden Eintrag des Dialogthemas, das heißt die Themenbezeichnung, in dem Dialogwissen.

```
1: Plan : {
2:   GOAL : PERFORM set-topic $topic;
3:   BODY :
4:     UPDATE (topic) (topic $topic);
5:     UTILITY : 50;
6: }
```



```
7: Plan : {
8:   GOAL : PERFORM get-topic $topic;
9:   BODY :
10:    ASSIGN $topic " ";
11:    WHEN : FACT topic $top {
12:      ASSIGN $topic $top;
13:    };
14:    UTILITY : 50;
15: }
```

6.3.2 Emuliertes Bewusstsein für Themeneinführungen

Um das künstliche Bewusstsein von Max hinsichtlich der **Angemessenheit** eines eingeführten Themas widerzuspiegeln, ist zunächst eine **Spezifikation adäquater Dialogthemen** für unterschiedliche Dialogsituationen und verschiedene Beziehungen zwischen den Dialogteilnehmern notwendig. Die Dialogsituation wird definiert durch den Kontext, in der der betrachtete Dialog stattfindet. Ausschlaggebend dabei ist die Umgebung oder auch der Grund für die Interaktion. Die Beziehung zwischen den Dialogteilnehmern hat ebenfalls Auswirkung auf die Angemessenheit von Themen. Daraus ergibt sich eine textuelle Spezifikation von Dialogthemen, wie sie in Tabelle 6.1 darge-

stellt ist. Die Information über die aktuell gegebene Dialogsituation und dem Bekanntheitsgrad, und somit die Auswahl der passenden Themenspezifikation, ist dem Agenten anhand eines entsprechenden Eintrags innerhalb seines Dialogwissens zugänglich.

Tabelle 6.1: Textbasierte Spezifikation adäquater Dialogthemen abhängig von Dialogsituation und Bekanntheitsgrad.

Dialogsituation	Bekanntheitsgrad	Adäquate Dialogthemen
Smalltalk	unbekannt	Wetter, Beruf, Hobbies, Sport, ...
	bekannt	wie unbekannt + Alter, Politik, ...
	befreundet	wie bekannt + Religion, Gesundheit, ...
Verkaufsgespräch	unbekannt	Mode, Hobbies, Sport, ...
	bekannt	wie unbekannt + Beruf, Familie, ...
...

Um nun verschiedene Reaktionen von Max auf die von dem menschlichen Dialogpartner vorgeschlagenen Themen auslösen zu können, enthält das Dialogwissen des Agenten eine Menge an die Themenspezifikation angepasste Regeln. Genauer gesagt überprüft der Dialogmanager des Dialogagenten einen Themenvorschlag des menschlichen Dialogpartners auf seine Adäquatheit durch einen Abgleich des **Satzthemas** der Äußerung (s. Abschnitt 5.3.4, Aufgabe 1) mit den in der Spezifikation enthaltenen Themen. Stimmen Bekanntheitsgrad und Angemessenheit des vorgeschlagenen Themas überein, so führt das System eine der bereits bestehenden Dialogregeln aus. Wenn nicht, so äußert der Agent sein Bewusstsein hinsichtlich der Unangemessenheit des Themas durch Empörung. Bei Äußerungen, für die sich kein Satzthema bestimmen lässt, deren Unangemessenheit jedoch sehr deutlich ist (beispielsweise bei der Frage „Wie alt bist du?“), führt eine entsprechende Interpretationsregel zu der Erkennung der Unangemessenheit. Die entsprechende Dialogregel zur Ausführung der Reaktion von Max ist JAM-Regel 4 zu entnehmen.

Darüber hinaus spiegelt sich die Angemessenheit von Themen auch bei den Themenvorschlägen seitens des Agenten wider. Das emulierte Bewusstsein für Themeneinführungen wird bei der bereits bestehenden Fähigkeit des Agentensystems, *proaktiv* ein neues Thema einzuführen, erreicht, indem die zugrunde liegenden Dialogregeln mit den Themen, die diese durch Ausführung vorschlagen, annotiert werden. Wird aus einem bestimmten Grund heraus systemintern das Ziel aufgeworfen, ein neues Thema anzusprechen, gelingt die Auswahl der Regel auch hier anhand eines Abgleichs der in

Algorithmus 4 JAM-Regel zur Ausführung einer passenden Reaktion von Max auf einen unangemessenen Themenvorschlag des menschlichen Dialogpartners. Durch Abfrage des Satzthemas der letzten Äußerung (Zeile 6) gelingt zudem die Benennung des unangemessenen Themas (Zeile 12). Konnte kein Satzthema ermittelt werden, äußert der Agent seine Empörung, ohne dabei auf das genaue Thema einzugehen (Zeile 8 – 10).

```

1: < rule name = "topics.reactToInadequateTopic" utility = "30" >
2: < match >
3: < convfunction type = "provide.topic.inadequate" / >
4: < /match >
5: < action >
6: < command function = "get-sentence-topic" arguments = "$topic" / >
7: < random >
8: < act > Ein sehr gewagtes Thema fuer ein erstes Gespraech! < /act >
9: < act > Das Thema $topic lassen wir im ersten Gespraech lieber! < /act >
10: < /random >
11: < /action >
12: < /rule >

```

den Regeln aufgezeichneten Annotationen mit der textbasierten Themenspezifikation.

6.3.3 Emuliertes Bewusstsein für Themenfortführungen

Ein Bewusstsein für die Fortführung eines Themas beinhaltet die Entscheidung darüber, ob ein vorgeschlagenes Thema tatsächlich zu dem gemeinsamen Thema des Dialogs gemacht werden soll. Zum vorliegenden Zeitpunkt wird die Fortführung eines Themas jedoch als **Ziel** des Agenten deklariert. Dies geschieht durch den Eintrag eines entsprechenden Beliefs innerhalb des Dialogwissens der Form

raise-dialogue-goal arguments = "system maintain-topic \$topic",

sobald ein Dialogthema *\$topic* identifiziert wurde.

6.3.4 Emuliertes Bewusstsein für Themenwechsel

Auf Grundlage der dritten Aufgabe des Themenerkenners wird der Dialogagent über etwa bevorstehende Themensprünge informiert. Handelt es sich dabei um eine von dem

menschlichen Dialogpartner initiierte Aktion, wird das Wissen darüber anhand einer entsprechenden Bemerkung des Agenten emuliert. Grundlage für diese Bemerkung liefert die in Algorithmus 5 aufgezeigte JAM-Regel.

Algorithmus 5 JAM-Regel zur Ausführung einer passenden Reaktion von Max auf einen von dem menschlichen Dialogpartner initiierten möglichen Themensprung. Dabei äußert der Agent entweder die Aufforderung, das alte Dialogthema weiter zu behandeln (Zeile 13), oder eine Verwunderung über den plötzlichen Themenwechsel (Zeile 14). Die Regel wird jedoch nur ausgeführt, wenn die Möglichkeit eines Themensprungs in dem Dialogwissen des Agenten eingetragen ist (Zeile 6) und der Agent das Ziel hat, das aktuelle Dialogthema weiter fortzuführen (Zeile 5).

```
1: < rule name = "user.topic-leap" utility = "20" >
2: < match >
3: < allof >
4: < convfunction type = "provide.topic.leap" / >
5: < testcheck = "goal" arguments = "maintain-topic" / >
6: < testcheck = "topic-leap" arguments = "Potential" / >
7: < /allof >
8: < /match >
9: < action >
10: < command function = "get-sentence-topic" arguments = "$stopic" / >
11: < command function = "get-topic" arguments = "$stopic" / >
12: < random >
13: < act > Lass uns doch erstmal beim Thema bleiben! < /act >
14: < act > Wie kommst du denn jetzt auf das Thema $stopic? < /act >
15: < /random >
16: < /action >
17: < /rule >
```

Die genannten Bemerkungen von Max als Reaktion auf einen möglichen Themensprung drücken in gewisser Weise eine Art Verwunderung oder Überraschung des Agenten gegenüber der plötzlichen Einführung eines komplett anderen Themas aus. In gewissen Situationen ist diese Art des Themenwechsels jedoch offensichtlich. Beispielsweise, wenn vor einem heranfahrenden Auto gewarnt wird. Abhängig von der jeweiligen Umgebung und der gegebenen Situation, in der sich der Dialog ereignet, ist daher eine Unterscheidung zwischen möglichen Themensprüngen, mit denen eine inhaltliche Änderung des weiteren Dialogverlaufs beabsichtigt wird, und solchen, die für beide Teilnehmer aufgrund der aktuellen Dialogsituation nachvollziehbar sind, zu treffen. Da das Agentensystem von Max nur begrenzt auf aktuelles Wissen über die

Umwelt, in dem der Agent situiert ist, zugreifen kann, ist eine automatische Umsetzung einer solchen Unterscheidung bisher nicht möglich.

Die Initiierung eines möglichen Themenwechsels kann aber auch proaktiv durch den Agenten selbst durchgeführt werden. Der Grund für die System-seitige Einleitung eines möglichen Themenwechsels ergibt sich aus der Dauer der aktuellen Dialogsequenz zu einem Thema. Zu diesem Zweck werden intern die zu einem Dialogthema geäußerten Dialogbeiträge gezählt. Übersteigt die Anzahl der gezählten Beiträge die Zahl 4 und wurde die letzte Äußerung des menschlichen Dialogpartners nicht als Frage interpretiert (die Beantwortung dieser hat gegenüber einer Themeneinführung Priorität) so veranlasst der Dialogmanager einen Themenvorschlag, der durch eine entsprechende Äußerung des Agenten durchgeführt wird. Realisiert wird dieser Vorgang erneut auf Grundlage einer Dialogregel, die eine Schablone für diese Äußerung bereithält (s. Algorithmus 6).

Algorithmus 6 JAM-Regel zur Initiierung eines möglichen Themensprungs. Ausgeführt wird diese Regel, wenn die Anzahl der Beiträge zu einem Dialogthema die Zahl 4 übersteigt (Zeile 3). In diesem Fall wird zunächst anhand des externen Programms „ChooseNewTopic“ ein neues Thema ausgewählt (Zeile 7) und dieses daraufhin durch eine Äußerung des Agenten vorgeschlagen (Zeile 8). Anschließend wird das Ziel aufgeworfen, eine entsprechende, das neue Thema einleitende Äußerung für den Agenten zu generieren (Zeile 9).

```
1: < rule name = "user.topic-change" utility = "20" >
2: < match >
3: < testcheck = "min-topic-length" arguments = "4" / >
4: < /match >
5: < action >
6: < block >
7: < command function = "ChooseNewTopic" arguments = "$new-topic" / >
8: < act > Lass uns doch ueber $new-topic reden. < /act >
9: < command function = "raise-dialogue-goal" arguments =
  system start-topic : $new-topic" / >
10: < /block >
11: < /action >
12: < /rule >
```

6.3.5 Emuliertes Bewusstsein für Themenwiedereinführung

Eine Themenwiedereinführung, die sich mit der folgenden Äußerung auf ein Themensprung ereignet, löst keine von den bisherigen Verhaltensweisen des Agenten abweichende Äußerung aus, um Max nicht doppelt auf einen Themensprung reagieren zu lassen. Eine Wiedereinführung eines bereits länger zurückliegenden Dialogthemas kommentiert der Agent hingegen mit der Äußerung „Über das Thema hatten wir doch schon gesprochen.“. Voraussetzung für die Erkennung einer solchen Themenwiedereinführung ist der Zugriff auf Informationen über bereits besprochene Dialogthemen. Zu diesem Zweck wird simultan zu dem laufenden Dialog eine **Themenhistorie** aufgebaut, die alle identifizierten Dialogthemen in textueller Form abspeichert. Mit jedem Themeneintrag in das Dialogwissen geht daher auch eine Aktualisierung der Themenhistorie einher. Dabei werden jedoch nur Dialogthemen hinzugefügt, die noch nicht in der Historie enthalten sind, um Redundanzen zu vermeiden. Die Wiedereinführung eines bereits abgeschlossenen Dialogthemas gelingt dann durch einen Stringabgleich des aktuellen Satzthemas mit der Themenhistorie. Dieser wird genau dann ausgeführt, wenn das System einen möglichen Themensprung identifiziert, der durch eine entsprechende Äußerung des menschlichen Gegenübers herbeigeführt wird.

Ein automatisch realisiertes Bewusstsein dafür, selbst, das heißt System-seitig, auf ein vergangenes Dialogthema zurückzukommen, lässt sich genau dann realisieren, wenn die Themenwiedereinführung anhand eines proaktiven Themenvorschlags des Agenten zustande kommt. Der Grund dafür liegt darin, dass die durch Regeln definierten Themenvorschläge thematisch annotiert sind (s. Abschnitt 6.3.2) und dem System daher bereits vor der Ausführung der Äußerung die Information hinsichtlich des damit eingeführten Themas zur Verfügung steht. Die Emulation des Bewusstseins für die Themenwiedereinführung im Rahmen eines Themenvorschlags gelingt dann durch die Äußerung „Lass uns doch nochmal über \$topic reden!“. Aktuell ist noch kein Auslöser für eine Themenwiedereinführung von Max realisiert.

6.4 Anwendungsbeispiel

Nachdem die Integration des Themenmodells in das Agentensystem von Max und insbesondere die Anpassung der Dialogregeln ausführlich vorgestellt wurden, beschreibt der vorliegende Abschnitt ein Anwendungsszenario, welches den Einsatz künstlichen

Themenbewusstsein in der natürlichen Mensch-Maschine-Interaktion reflektiert. Es beginnt mit der Darstellung der betrachteten Dialogsituation zwischen dem Agenten Max und seinem menschlichen Dialogpartner und erläutert darauf aufbauend die Spezifikation einer Menge an möglichen Dialogthemen für den automatischen Themenkenner. Daraufhin wird durch die Präsentation eines Beispieldialogs ein Einblick in den finalen Dialogverlauf zwischen Mensch und Agent gegeben.

6.4.1 Dialogszenario

Die grundlegende Gesprächsart bei der Anwendung des Agenten Max als Rezeptionisten und Museumsführer bildet der **Smalltalk**. Der Begriff „Smalltalk“ wird hier aufgefasst als alltägliche Unterhaltung zwischen Unbekannten, aber auch zwischen Bekannten (s. *casual conversation*, (Ventola, 1979)), die nicht über oberflächliche Inhalte hinausgeht. Dementsprechend handelt es sich bei der ersten realisierten Anwendung um einen typischen Smalltalk-Dialog, der sich hinsichtlich zwei unterschiedlichen Bekanntheitsgraden unterscheiden kann. Die textbasierte Themenspezifikation für das umgesetzte Beispielszenario definiert daher zum einen Themen, die mögliche Dialogthemen zwischen Unbekannten und zum anderen mögliche Dialogthemen zwischen Bekannten präsentieren. Um einen kohärenten Dialog zwischen Max und seinem menschlichen Dialogpartner zu ermöglichen, wurde zudem eine Vorgabe hinsichtlich der thematischen Struktur des Dialogs ausgearbeitet, welche die Einbettung des künstlichen Themenbewusstseins in den Dialog ermöglicht und ein entsprechendes Konversationsverhalten des Agenten hervorruft. Dialogsituation und Dialogstruktur werden im Folgenden genauer beschrieben.

Dialogsituation und Dialogstruktur

Die betrachtete **Dialogsituation** für das Anwendungsbeispiel entspricht dem Einsatz des Agenten Max als Rezeptionisten auf dem Flur der Arbeitsgruppe Wissensbasierte Systeme in dem Gebäude der Universität Bielefeld. Da viele Besucher der Arbeitsgruppe sowie Studenten, die zufällig den Flur entlang gehen, mit Max interagieren, kommt es häufig zu Smalltalk-Situationen, bei denen sich die beteiligten Dialogpartner zuvor nicht kennen. In Interaktionen mit Mitarbeitern und wiederkehrenden Gästen sind sich künstlicher und menschlicher Dialogpartner hingegen bereits bekannt.

Mit der Festlegung der Dialogsituation auf Smalltalks geht die Umsetzung einer dieser Dialogform zugrunde liegenden Struktur einher. Zu diesem Zweck sind die das konversationale Verhalten des Agenten definierenden Dialogregeln für das hier betrachtete Szenario entsprechend der von Schneider (1988) festgelegten **Dialogstruktur typischer Smalltalk-Situationen** aufgebaut. Diese besitzt im Allgemeinen folgenden Aufbau:

1. Frage
2. Antwort
3. Rückantwort
4. Weitere Beitragswechsel

Dieser Struktur folgend, stellt der Agent seinem menschlichen Gegenüber zur Einleitung eines neuen Dialogabschnitts eine Frage, auf die der Agent eine Antwort erwartet. Je nachdem, ob diese Antwort Raum für weitere Fragen lässt und ob das Agentensystem über weitere Dialogregeln zu diesem von dem Agenten angesprochenen Thema verfügt, vertieft Max den Dialoginhalt an dieser Stelle durch eine weitere Frage oder Anmerkung. Auf diese Weise kommt eine Interaktion zustande, welche die Grundlage für die Erkennung der unterschiedlichen Aktionen, die den Ablauf eines thematischen Dialogs beeinflussen, liefert und in der das künstliche Bewusstsein von Max für diese Aktionen zu den gewünschten konversationalen Verhaltensweisen des Agenten führen kann. Darüber hinaus liefert diese Dialogstruktur die Grundlage für die Abfrage und Analyse der Interessen der einzelnen Dialogpartner des Agenten, welche zukünftig dem Aufbau eines *Personengedächtnisses* von Max dienen sollen (s. hierzu (Mattar & Wachsmuth, 2012)).

Festlegung möglicher Dialogthemen

Unterschiedlichen Dialogszenarien liegen unterschiedliche Mengen an möglichen Dialogthemen zugrunde. Demnach ist die Definition einer entsprechenden **Themenspezifikation** erforderlich, die dem System wie in Abschnitt 6.3.2 beschrieben in Form einer Textdatei zugänglich gemacht wird. Die Wahl der darin spezifizierten Dialogthemen hängt dabei von der jeweiligen Dialogsituation ab und sollte sogenannte *Tabuthemen*

spezifizieren, die für ein erstes Zusammentreffen der Dialogteilnehmer zu persönlich und somit unangemessen sind.

Die im Rahmen des *CUBE-G*-Projekts (M. Rehm et al., 2008) zum Aufbau eines standardisierten Korpus multimodaler Interaktionen zwischen zwei Personen durchgeführten und dokumentierten Kennenlerngespräche fanden ebenfalls in einem universitären Umfeld statt. Die dabei zusammengetragenen Themen liefern aufgrund der ähnlichen Dialogsituation eine geeignete Grundlage für die Spezifikation angemessener Dialogthemen für die Flurinstallation des Agenten Max. Diese Menge wurde um einige Dialogthemen erweitert, um einerseits eine flexiblere Themenerkennung zu gewährleisten und andererseits eine gewisse Anzahl an Themen definieren zu können, die in einem ersten Smalltalk tabu, in weiteren Dialogen mit derselben Person jedoch als angemessen empfunden werden. Tabelle 6.2 liefert die daraus resultierende Liste der möglichen Dialogthemen. Zu jedem dieser Themen ist zudem die Wikipedia-Kategorie angegeben, welche das entsprechende Dialogthema repräsentiert und auf dessen Grundlage die Themenerkennung realisiert wird. Themen, die bereits in einem ersten Smalltalk zwischen zwei Personen als angemessen empfunden werden, sind oberhalb der gestrichelten Linie aufgeführt, solche, die in einer derartigen Smalltalk-Situation nach Schneider (1988) als tabu deklariert werden, unterhalb.

6.4.2 Beispieldialog

Mit dem Ziel, die Ausführung adäquater Äußerungen des Agenten hinsichtlich der aktuellen Themensituation zu veranschaulichen, liefert Abbildung 6.5 die Darstellung eines Beispieldialogs zwischen einer Person (P) und Max (M). Die zugrunde liegende Dialogstruktur und -situation entsprechen dabei den zuvor vorgenommenen Spezifikationen. Bei Betrachtung der einzelnen Äußerungen des Dialogs ist die vorgegebene Frage-Antwort-Struktur klar erkennbar. So führt Max den Dialogzugang, also den thematischen Teil des Smalltalks, mit einer Frage nach der Herkunft seines Gegenübers ein („*Woher kommst du denn?*“). Ausgehend von der Antwort auf diese Frage zieht Max anschließend Rückschlüsse über den Lieblingsfußballverein seines Dialogpartners und stellt diesbezüglich eine weitere Frage („*Bist du denn Fan von Armina Bielefeld?*“). Im weiteren Dialogverlauf entwickelt sich zunächst ein kohärentes Gespräch über Fußball, welches der Agent nach einigen Dialogbeiträgen durch Vorschlag eines anderen Themas zu beenden versucht („*Lass uns doch mal über Bildung reden.*“). Die Person

Tabelle 6.2: Liste der vordefinierten Dialogthemen für das betrachtete Beispielszenario mit Max und der Wikipedia-Kategorien, die diese Dialogthemen repräsentieren. Die Angaben unterhalb der gestrichelten Linie entsprechen Dialogthemen, die erst mit zunehmendem Bekanntheitsgrad der Dialogteilnehmer als angemessen empfunden werden und für den ersten Smalltalk der Dialogpartner somit tabu sind.

Dialogthema	Wikipedia-Kategorie
Familie	Familie
Freizeit	Freizeit
Ort	Region als Thema
Beruf	Beruf
Gesundheit	Gesundheit
Musik	Musik
Firmen	Unternehmen
Literatur	Literatur
Bildung	Bildung
Medien	Medien
Tiere	Vielzellige Tiere
Sport	Sport
Fußball	Fußball
Basketball	Basketball
Essen und Trinken	Essen und Trinken
Wetter	Meteorologie
Natur	Umwelt und Natur
Geschichte	Geschichte
Autos	Automobil
Politik	Politik
Religion	Religion
Tod	Tod
Sex	Sexualität des Menschen

geht auf diesen Vorschlag ein, sodass das Thema Bildung in weiteren Dialogbeiträgen fortgeführt wird. Der Dialog endet mit einer Verabschiedung.

Die genauere Betrachtung der zur Laufzeit des Dialogs ermittelten Themeninformationen und der zugrunde liegenden Dialogregeln gibt Aufschluss über die Generierung der einzelnen Äußerungen von Max. Mit der Identifikation des ersten Dialogthemas (Ort) wird intern das Ziel aufgeworfen, dieses Thema weiter aufrecht zu halten. Die Äußerung der Person „*Wie kommt man denn von hier zur Mensa?*“ signalisiert dem System einen möglichen Themensprung, den der Agent aufgrund seines aufgeworfenen Ziels zu verhindern versucht. Anhand der entsprechenden, innerhalb des Programmco-

	Utterance	Dialog Topic	Topic Leap?
P:	Hallo	---	No
M:	Hi ich bin Max! Wie heisst Du?	---	No
P:	Stefan	---	No
M:	Hallo Stefan ! Freut mich dich kennenzulernen. Woher kommst du denn?	---	No
P:	Aus Bielefeld	Ort (0.43)	No
M:	Das ist ja cool, ich bin auch Bielefelder	Ort (0.79)	No
P:	wie kommt man denn von hier zur Mensa?	---	Potential
M:	Lass uns doch erstmal beim Thema bleiben. Bist du denn Fan von Arminia Bielefeld?	Geschichte (0.41)	No
P:	Ja ich bin ein grosser Arminia Fan	Fußball (0.44)	No
M:	Ja super, ich finde Arminia auch toll	Fußball (0.53)	No
P:	Schade dass sie letztes Mal gegen Hansa Rostock verloren haben	Fußball (0.54)	No
M:	Ja ich bin immer traurig wenn Arminia verliert	Fußball (0.54)	No
P:	Naja so schlimm ist es bei mir nicht	Fußball (0.54)	No
M:	Lass uns doch mal über Bildung reden. Was studierst du?	Literatur (0.36)	Potential
P:	Ich studiere Informatik an der Uni hier	Bildung (0.54)	Yes
M:	Dann kennst du bestimmt Ipke Wachsmuth	Bildung (0.48)	No
P:	Den kenn ich. Ich muss jetzt aber mal die Mensa suchen. Bis bald	Bildung (0.48)	No
M:	Tschuess Stefan - bis zum naechsten mal.	---	No

Abbildung 6.5: Beispieldialog zwischen Max (M) und einer Person (P).

des 5 aufgeführten Dialogregel fordert Max seinen Dialogpartner daher auf, beim alten Thema zu bleiben. Mit der direkt an dieser Äußerung anschließenden Frage „*Bist du denn Fan von Arminia Bielefeld?*“ führt Max zwar nicht das alte Thema „Ort“ weiter fort, bewahrt jedoch die Kohärenz des Dialogs, sodass der mögliche Themensprung letztendlich nicht vollzogen wird (s. Spalte „Topic Leap“).

Der weitere Dialog zum Thema Fußball verdeutlicht, dass sich Max bezüglich des Dialogthemas mit weiteren Äußerungen „sicherer“ wird und die Werte für dieses Thema entsprechend steigen. Die dann von Max initiierte Möglichkeit eines Themensprungs ist auf die in dem Programmcode 6 dargestellte Dialogregel zurückzuführen. Demnach wurden mehr als 4 Beiträge zu demselben Thema beigetragen, was zu der Ausführung eines Themenvorschlags seitens des Agenten führte. Die Tatsache, dass der mensch-

liche Dialogpartner auf den Themenvorschlag von Max eingeht und das Thema Bildung mit seiner anschließenden Äußerung fortführt, bewirkt die Identifikation dieses Themas als aktuelles Dialogthema und zudem die Erkennung eines durchgeführten Themensprungs (s. „Yes“ in der Topic Leap-Spalte). Auf Basis der Verabschiedung der Dialogpartner, wie auch auf Basis ihrer Begrüßung, wird kein Dialogthema identifiziert, da diese Dialogabschnitte gemäß Ventola (1979) keinen thematischen Inhalt besitzen.

Die einzige Schwachstelle innerhalb dieses Beispieldialogs betrifft die Identifikation eines falschen Dialogthemas. So wird zum Zeitpunkt eines möglichen Themensprungs zwar kein Dialogthema identifiziert, die in der Äußerung „*Wie kommt man denn von hier zur Mensa?*“ genannten Konzepte werden an dieser Stelle jedoch in den weiteren Erkennungsprozess einbezogen. Dies bedingt anschließend die Identifikation des falschen Dialogthemas „Geschichte“. Wie zuvor beschrieben, werden die Konzepte einer möglichen Themensprung einleitenden Zwischenbemerkung im Allgemeinen nicht weiter bei der Themenerkennung beachtet. Voraussetzung dafür ist jedoch die darauffolgende Erkennung einer Themenwiedereinführung. Im vorliegenden Fall wird mit der nächsten Äußerung ein Themendrift vollzogen, sodass die zuvor genannten Konzepte, in diesem Fall das Konzept „Mensa“, doch in die Identifikation des Dialogthemas einbezogen werden. Dies führt dann letztendlich dazu, dass das Thema „Geschichte“ mit Verarbeitung der nächsten Äußerung als aktuelles Dialogthema ermittelt wird.

6.5 Zusammenfassung

Nachdem in Kapitel 5 die technische Umsetzung eines automatischen Themenerkenners zur Ermittlung thematischer Informationen dargestellt wurde, beschäftigte sich dieses Kapitel mit der Beschreibung der daran anschließenden Integration des resultierenden Themenwissens in das Konversationsverhalten des virtuellen Agenten Max. Der erste Abschnitt dieses Kapitels vermittelte die dafür erforderlichen Informationen hinsichtlich der bestehenden Dialogfähigkeit des Agenten. Dabei wurde herausgestellt, dass sowohl die Interpretation der sprachlichen Eingaben des menschlichen Dialogpartners als auch die Generierung adäquater Äußerungen von Max auf der Ausführung von Dialogregeln beruhen.

Im Anschluss daran erfolgte die Darstellung der physikalischen Einbettung des automatischen Themenerkenners in die Gesamtarchitektur des Agenten. Letztere basiert auf einem Multiagentensystem, sodass der erste Integrationsschritt die Implementierung eines Themenagenten war, dessen Aufgabe in der Durchführung der Erkennungsprozesse besteht. Darauf aufbauend wurde die Weiterleitung der mit jeder Äußerung des betrachteten Dialogs aktualisierten Themeninformationen an die das konversationale Verhalten von Max realisierende Komponente, den Dialogagenten, beschrieben. Diese bildete die Grundlage für eine Integration der Informationen in das Dialogwissen des Agenten.

Der nachfolgende Abschnitt befasste sich mit der Herausforderung, das hinzugewonnene Themenwissen in die regelbasierte Generierung adäquater Äußerungen für Max einfließen zu lassen. Für jede Aktion, die den thematischen Ablauf eines Dialogs beeinflussen, wurden neue Dialogregeln vorgestellt, die das künstliche Bewusstsein von Max anhand entsprechender Verhaltensweisen widerspiegeln. Die Darstellung der dafür erforderlichen Schritte erfolgte separat für jede einzelne Aktion.

Das durch die beschriebenen Integrationsschritte erlangte künstliche Themenbewusstsein des Agenten Max wurde schließlich durch die Darstellung eines konkreten Anwendungsbeispiels veranschaulicht. Dabei wurde zunächst eine Beschreibung des Dialogszenarios geliefert, welches sich aus der zugrunde liegenden Dialogsituation, der vorgegebenen Dialogstruktur und den möglichen Dialogthemen zusammensetzt. Daran anschließend wurde ein Beispieldialog zwischen Max und einer Person geliefert. Der daraus resultierende Verlauf des Dialogs ließ einen natürlichen, kohärenten Ablauf erkennen.

7 Resümee

Zum Abschluss soll rückblickend auf die gesamte Ausarbeitung ein Resümee gezogen werden. Zu diesem Zweck wird zunächst eine Zusammenfassung der in der Arbeit zusammengetragenen Ergebnisse geliefert. Im Anschluss daran wird die Frage diskutiert, ob das in der Einleitung formulierte Ziel tatsächlich erreicht und damit einhergehend ein Beitrag zur Verbesserung der Interaktion zwischen Menschen und sprachbasierten Systemen geleistet wurde. Des Weiteren werden im Ausblick Überlegungen hinsichtlich weiterführender Arbeiten angestellt, die ergänzende, innerhalb der vorliegenden Arbeit bislang ausgeklammerte Aspekte künstlichen Themenbewusstseins aufgreifen.

7.1 Ergebnisse

Ausgehend von der Motivation, eine intuitive und dadurch dem Menschen angepasste Interaktion mit Maschinen zu ermöglichen, lag das Ziel der vorliegenden Arbeit in der Realisierung künstlichen Themenbewusstseins in natürlichen Dialogen zwischen einem menschlichen und einem virtuellen Dialogpartner. Anhand dieser Zielsetzung ließen sich zwei Aufgabenschwerpunkte für die Arbeit formulieren. So bestand die erste Aufgabe in der Konzeption und Umsetzung eines Themenmodells, welches die automatische Erkennung und Analyse von Gesprächsthemen in natürlichen Dialogen ermöglicht. Dabei galt es den Forderungen nach Generalität und Echtzeitfähigkeit des Themenmodells nachzukommen. Die zweite Aufgabe umfasste die Integration der aus der Themenerkennung resultierenden Themeninformationen in das Dialogverhalten des konversationalen Agenten Max. Bei der damit verbundenen Einbettung des Themenmodells in die Gesamtarchitektur musste die bestehende Echtzeitfähigkeit des Systems bewahrt werden. Die anschließende Erweiterung des Dialogverhaltens des Agenten musste der Forderung nach Natürlichkeit gerecht werden.

Um ein genaues Verständnis der Eigenschaften menschlicher Dialoge sowie der menschlichen Fähigkeit, sich eines Gesprächsthemas bewusst zu sein, zu vermitteln, wurden

in Kapitel 2 die entsprechenden Theorien hinsichtlich dieses Forschungsgebiets zusammengetragen und grundlegende Begriffe festgelegt. Hierbei wurden insbesondere der Dialog als kommunikative Handlung zwischen zwei Personen und das Thema als Reflexion des Gesprächsinhalts definiert. Beiden Konzepten wurde darüber hinaus die Betrachtungsweise als Produkt der wechselseitigen Zusammenarbeit der beteiligten Dialogpartner zugesprochen. Des Weiteren wurde das Themenbewusstsein als Fähigkeit, Gesprächsthemen sowohl wahrzunehmen als auch in das individuelle Gesprächsverhalten einzubeziehen, spezifiziert. Auf Grundlage dieses Bewusstseins sind Menschen dazu in der Lage, den thematischen Ablauf eines Gesprächs durch bewusste Aktionen zu beeinflussen. Diese Aktionen wurden spezifiziert als die Einführung, Fortführung, Wiedereinführung und der Wechsel von Themen. Schließlich wurde das mentale Lexikon als Quelle menschlichen Wissens über Konzepte und ihre Zusammenhänge vorgestellt und als grundlegende Voraussetzung für den Umgang mit Gesprächsthemen identifiziert.

Mit dem Ziel, Anregungen für die technische Realisierung der in Kapitel 2 gesammelten Erkenntnisse hinsichtlich menschlichen Themenbewusstseins innerhalb von Mensch-Maschine-Schnittstellen zu erlangen, lieferte Kapitel 3 einen Überblick über entsprechende Ansätze. Dabei wurde in einem ersten Schritt die deutsche Wikipedia als Ressource für maschinenzugängliches Konzeptwissen und somit als künstliches mentales Lexikon für die automatische Themenerkennung festgelegt. Ausgehend von der Tatsache, dass das in der Wikipedia enthaltene Wissen zwar teilstrukturiert, die Inhalte jedoch größtenteils in Form von natürlichsprachlichen Texten vorliegen, wurden anschließend texttechnologische Verfahren zur Berechnung von Textähnlichkeiten eingeführt. Der Fokus wurde auf das Vektorraummodell gelegt, eines der verbreitetsten Modelle, auf dessen Basis die Wikipedia-Inhalte in Form von Vektoren dargestellt und somit maschinell vergleichbar werden. Die sich daran anschließende Problematik der automatischen Textkategorisierung und somit der Bestimmung eines gemeinsamen Themas der Texte bedurfte nur weniger Diskussion, da die Wikipedia-Inhalte bereits in Form eines Kategoriengraphen thematisch sortiert vorliegen. In diesem Zusammenhang wurden bestehende Ansätze für die automatische Kategorisierung von Texten auf Basis dieses Graphen vorgestellt. Als Fazit der Darstellung wurde die Idee, alle in dem jeweiligen Textdokument enthaltenen Konzepte auf Wikipedia-Artikel abzubilden und anschließend die gemeinsamen Kategorien dieser Artikel als Thema des Dokuments zu

definieren, als Ansatz für unsere automatische Themenerkennung übernommen. Anschließend wurde das TDT-Projekt vorgestellt, was zu der Ableitung einer weiteren Konzeptentscheidung, nämlich der Unterteilung des Erkennungsprozesses in einzelne Aufgaben, führte. Zum Schluss wurde ein Überblick über den aktuellen Stand der Forschung hinsichtlich künstlichen Themenbewusstseins in konversationalen Systemen geliefert. Dieser verdeutlichte die Notwendigkeit eines neuen Ansatzes für den Prozess der automatischen Themenerkennung, welcher die bestehenden Nachteile wie der hohe manuelle Aufwand, die Einschränkung auf bestimmte Domänen und die Vernachlässigung des interaktiven Aufbaus eines Themas behebt.

Auf Basis der in den Kapiteln 2 und 3 gesammelten Erkenntnisse und der zu Beginn der Arbeit aufgezeigten Aufgabenbeschreibung wurde im darauffolgenden Kapitel 4 die Konzeption der Gesamtarchitektur für die Realisierung künstlichen Themenbewusstseins für den Agenten Max vorgenommen. Die Einteilung dieser Architektur in zwei Bestandteile zur Bewältigung der beiden zuvor definierten Aufgabenschwerpunkte führte dabei zu der Entwicklung von zwei separaten, aufeinander aufbauenden Konzepten. Die Schnittstelle bilden dabei die Themeninformationen, die von dem ersten Bestandteil, dem Themenmodell, zur Verfügung gestellt und innerhalb des zweiten Bestandteils auf den konversationalen Agenten Max übertragen werden. Für das Themenmodell wurde der automatische Themenerkennung als Hauptkomponente konzipiert, dessen Funktion in der Identifikation des aktuellen Dialogthemas besteht. Dafür wurde zunächst eine präzise Definition eines Dialogthemas festgelegt. Aufbauend auf dieser Definition und den Forderungen nach Generalität und Echtzeitfähigkeit wurde eine Architektur für einen dynamischen, online-fähigen und globalen Themenerkennung entwickelt. Daraufhin wurden weitere, während des eigentlichen Prozess zur Themenerkennung auszuführende Aufgaben für den Themenerkennung formuliert, welche auf die Ermittlung zusätzlicher Informationen über die aktuelle Dialogsituation abzielen. Jede Information entspricht dabei einer Auskunft bezüglich der in Kapitel 2 aufgezählten Aktionen der Dialogteilnehmer, die den thematischen Verlauf eines Gesprächs beeinflussen.

Für den zweiten Bestandteil der Gesamtarchitektur, die Integration der ermittelten Themeninformationen, wurden zwei Teilschritte festgelegt. Der erste Teilschritt umfasste die physikalische Einbettung des Themenerkenners in die bestehende Systemarchitektur des Agenten Max, sodass die bestehende Echtzeitfähigkeit des Agenten

gewährleistet bleibt. Als Lösung wurde die Weiterleitung der Themeninformationen an die Deliberative Komponente, welche das konversationale Verhalten von Max realisiert, über eine Schnittstelle festgehalten. Der Konzeption des zweiten Teilschritts, der themenbasierten Anpassung des Konversationsverhaltens, lag die Forderung nach Natürlichkeit zugrunde. Demnach wurden Verhaltensweisen festgelegt, die das Bewusstsein des Agenten hinsichtlich der thematischen Aktionen durch die Generierung adäquater Äußerungen widerspiegeln.

In Kapitel 5 wurde schließlich die technische Umsetzung des in Kapitel 4 konzipierten Themenmodells beschrieben. Der Schwerpunkt lag dabei auf der Realisierung des Themenerkenners, deren Ergebnis die Entwicklung eines Algorithmus zur automatischen Erkennung der Themeninformationen in laufenden Gesprächen war. Die Durchführung des Algorithmus mit jedem neuen Dialogbeitrag gewährleistet die Aktualität der resultierenden Informationen. Durch Abbildung der in den einzelnen Äußerungen enthaltenen Konzeptterme auf Wikipedia-Kategorien gelingt hierbei die Erkennung möglicher thematischer Richtungen einer Äußerung. Durch den anschließenden Vergleich dieser Richtungen mit der nachfolgenden Äußerung können thematische Überschneidungen und darauf aufbauend das gemeinsame Thema dieser Äußerungen identifiziert werden. Dabei lassen sich simultan mit der Ausführung der einzelnen Erkennungsprozesse die weiteren Themeninformationen ableiten. Die Evaluation des Themenerkenners verdeutlichte dessen erfolgreiche Umsetzung, aber auch die Schwierigkeit, Themen auf der Grundlage einzelner Äußerungen zu erkennen.

Gegenstand des sechsten Kapitels war die Übertragung der aus der automatischen Themenerkennung hervorgehenden Themeninformationen auf den Agenten Max. Aufgrund der Tatsache, dass die bestehende Architektur von Max auf einem Multiagentensystem aufbaut, wurde für die physikalische Integration des Themenerkenners ein Themenagent implementiert. Die Kommunikation der von dem Themenerkennner ermittelten Themeninformationen wurde durch die Übersendung dieser Informationen in Form einer entsprechenden Nachricht des Themenagenten an den Dialogagenten sichergestellt. Letzterer verarbeitet die Eingaben des menschlichen Dialogpartners von Max auf der Grundlage von Dialogregeln, welche die Ausführung einzelner Pläne für die Generierung entsprechender Antworten des konversationalen Agenten anstoßen. Um auf Basis der aktuellen Informationen das konversationale Verhalten von Max zu beeinflussen, wurden daher neue Dialogregeln definiert, die, abhängig von der jeweili-

gen Dialogsituation, die Ausführung adäquater Äußerungen auslösen. Auf diese Weise gelang schlussendlich die Realisierung künstlichen Themenbewusstseins in der natürlichen Interaktion zwischen Max und einem menschlichen Dialogpartner, was anhand eines Beispieldialogs belegt werden konnte.

7.2 Fazit

Auf Grundlage der finalen Darstellung der erbrachten Ergebnisse lässt sich abschließend ein Fazit hinsichtlich der Frage ziehen, inwiefern die hier beschriebenen Lösungsansätze einen Beitrag zu der Erreichung des zu Beginn der Arbeit formulierten Ziels leisten. Das Ziel war dabei die Entwicklung eines künstlichen Themenbewusstseins, welches die Erkennung von Gesprächsthemen in natürlichen Dialogen und die Integration der daraus resultierenden Themeninformationen in die Dialogkomponente des konversationalen Agenten Max beinhaltet. Die Konzeption und Implementation waren dabei so durchzuführen, dass die Emulation menschlichen Themenbewusstseins gelingt und die kommunikativen Fähigkeiten des Agenten mittels dieses Bewusstseins erweitert werden.

In Bezug auf die automatische Themenerkennung ist es uns gelungen, einen Ansatz für ein Themenmodell auszuarbeiten und umzusetzen, welches nicht nur zu jedem Zeitpunkt der betrachteten Interaktion das aktuelle Dialogthema, sondern darüber hinaus weitere Informationen über die thematischen Abläufe innerhalb des Dialogs bereitstellt. Auf diese Weise wird eine ganzheitliche Betrachtung der aktuellen Themensituation des Dialogs und darauf aufbauend die automatische Interpretation von sprachlichen, den weiteren thematischen Dialogverlauf beeinflussenden Aktionen der teilnehmenden Sprecher ermöglicht. Das Fundament für die Umsetzung dieser umfangreichen Themenerkennung lieferte die Einbindung der theoretischen Grundlagen hinsichtlich menschlichen Themenbewusstseins und die Abtrennung der zugrunde liegenden Prozesse von der bestehenden regelbasierten Eingabenverarbeitung des Agentensystems. So wurde ein Modell ausgearbeitet, welches die bestehenden Ansätze zur Erkennung von Dialogthemen hinsichtlich der Vollständigkeit der ermittelten Informationen und der Berücksichtigung fundierter Theorien erweitert.

Die anschließende Integration der Themeninformationen führte schließlich zu der angestrebten Realisierung künstlichen Themenbewusstseins für den Agenten Max.

Durch eine echtzeitfähige Einbindung des automatischen Themenerkenners und der Definition von Dialogregeln, die das dadurch erlangte Themenwissen in das Konversationsverhalten von Max einbeziehen, gelang zudem die gewünschte Erweiterung der konversationalen Fähigkeiten von Max. So generiert das zugrunde liegende System des Agenten Äußerungen, welche das künstliche Bewusstsein von Max über die aktuelle Themensituation reflektieren.

Anhand der Evaluationsergebnisse bezüglich des automatischen Themenerkenners und des in dem Kapitel 6 dieser Ausarbeitung dargestellten Beispieldialogs ließen sich einige Schwachstellen der hier vorgestellten Lösungen ausmachen. So sind weiterführende Arbeiten, die die Akkuratheit des Themenerkenners erhöhen und das Konversationsverhalten von Max weiter ausbauen, durchzuführen, um die erkannten Schwachstellen gezielt zu beheben. Der nachfolgende Abschnitt liefert diesbezüglich einen Ausblick.

Insgesamt wurde mit der vorliegenden Arbeit eine Grundlage für die Emulation menschenähnlichen Themenbewusstseins in konversationalen Agenten geschaffen. Der Ansatz, eine sprachliche Schnittstelle mit künstlichem Themenbewusstsein auszustatten, um den natürlichen Dialog zwischen Mensch und Maschine zu verbessern, stellt dabei eine neue Überlegung dar, die in bisherigen Arbeiten noch nicht betrachtet wurde. Die damit verbundenen Lösungsansätze leisten somit nicht nur einen Beitrag zur Erreichung des genannten Ziels, sondern liefern darüber hinaus neue Erkenntnisse im Bereich der natürlichen Mensch-Maschine-Interaktion.

7.3 Ausblick

Bislang durchgeführte Studien bezogen sich ausschließlich auf die Evaluation der automatischen Themenidentifikation. Ein wichtiger nächster Schritt wäre daher zum einen die Evaluation der von dem Themenerkennner zusätzlich bereit gestellten Informationen und zum anderen die Durchführung einer Nutzerstudie, welche die Qualität des künstlichen Themenbewusstseins in der direkten Interaktion mit Max untersucht. Letztere bietet zudem die Möglichkeit, weitere Erkenntnisse bezüglich der Wahrnehmung und Glaubwürdigkeit von konversationalen Agenten zu erlangen. Aufbauend auf den durch diesen Evaluationen erlangten Ergebnissen können dann weitere Optimierungsschritte geplant und umgesetzt werden.

Neben der Optimierung der Themenerkennung und der Durchführung zusätzlicher Evaluationen, sollten zukünftige Arbeiten darüber hinaus auf die Erweiterung des bestehenden Ansatzes abzielen. Einen wesentlichen Aspekt hierfür liefert die taxonomische Anordnung des Kategoriengraphen der Wikipedia. So bietet diese Struktur die Möglichkeit, für die vordefinierten Kategorien, welche im Rahmen der vorliegenden Arbeit die Menge an möglichen Dialogthemen repräsentieren, anhand der *Hyperonymie (isa)*-Beziehungen sowohl Ober- als auch Unterkategorien zu ermitteln und darauf aufbauend Ober- und Unterthemen zu identifizieren. So lässt sich zusätzlich ein Bewusstsein für den Zusammenhang zwischen Dialogthemen und möglichen spezielleren Unterthemen emulieren.

Ein aufgrund des gewählten Szenarios bisher gänzlich ausgeklammerter Gesichtspunkt stellt die Betrachtung zusätzlicher Hinweise auf die Durchführung thematischer Aktionen dar, welche sich vorwiegend im Rahmen gesprochener Dialoge ereignen. Gemeint sind hiermit beispielsweise formulierte Wiederholungen, Zusammenfassungen und Schlussfolgerungen der Sprecher, die auf die Beendigung des aktuellen Themas hindeuten (Howe, 1991; Svennevig, 1999). Weitere Anhaltspunkte liefern die Äußerung bestimmter Wörter der Dialogteilnehmer oder längere Gesprächspausen. Voraussetzung für die Einbindung dieser Faktoren ist jedoch ein auf gesprochener Sprache basierender Dialog zwischen Mensch und konversationaler Schnittstelle. Genau wie die Realisierung künstlichen Themenbewusstseins bildet dieser Aspekt einen weiteren Beitrag, die im Eingangszitat von Schank (1977) angesprochene anspruchsvolle Problematik „Kommunikation“ technisch zu bewältigen.

Literaturverzeichnis

- Adams, P. & Martel, C. (2010). Conversational thread extraction and topic detection in text-based chat. In P. C.-Y. Sheu, H. Yu, C. V. Ramamoorthy, A. K. Joshi & L. A. Zadeh (Hrsg.), *Semantic Computing*. IEEE Press.
- Aitchison, J. (2003). *Words in the Mind: An Introduction to the Mental Lexicon* (3. Aufl.). Blackwell.
- Allan, J. (2002). Introduction to topic detection and tracking. In J. Allan (Hrsg.), *Topic Detection and Tracking: Event-based Information Organization* (S. 1-16). Kluwer Academic Publishers.
- Altmann, G. T. (1997). *The Ascent of Babel: an Exploration of Language, Mind, and Understanding*. Oxford University Press.
- Aronoff, M. (2003). *The Handbook of Linguistics* (M. Aronoff & J. Rees-Miller, Hrsg.). Blackwell.
- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. ACM Press (u.a.).
- Bazzanella, C. & Dmiano, R. (1997). Coherence and misunderstanding in everyday conversations. In W. Bublitz, U. Lenk & E. Ventola (Hrsg.), *Coherence in Spoken and Written Discourse – How to create it and how to describe it*. John Benjamins Publishing.
- Bergstrom, T. & Karahalios, K. (2009). Conversation clusters: Grouping conversation topics through human-computer dialog. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)* (S. 2349-2352). Boston, USA.
- Bernsen, N. O. & Dybkjær, L. (2005). Meet Hans Christian Andersen. In *Proceedings of the Sixth SIGdial Workshop on Discourse and Dialogue* (S. 237-241). Lisbon, Portugal.
- Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R. et al. (2009). DBpedia – A crystallization point for the web of data. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 7, 154–165.
- Boone, G. (1998). Concept features in Re:Agent, an intelligent email agent. In *Proceedings of the Second International Conference on Autonomous Agents (AGENTS '98)* (S. 141–148). Minneapolis, USA.

- Breuing, A. & Wachsmuth, I. (2010a). Equipping a conversational agent with access to Wikipedia knowledge. In *Proceedings of the KogWis 2010: 10th Biannual Meeting of the German Society for Cognitive Science* (S. 89). Potsdam, Germany.
- Breuing, A. & Wachsmuth, I. (2010b). Improving human-agent conversations by accessing contextual knowledge from Wikipedia. In *Proceedings of the 3rd WI-IAT Doctoral Workshop, in conj. with the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2010)* (S. 428-431). Toronto, Canada.
- Breuing, A. & Wachsmuth, I. (2012). Let's talk topically with artificial agents! Providing agents with humanlike topic awareness in everyday dialog situations. In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART 2012)* (S. 62-71). Vilamoura, Portugal.
- Breuing, A., Waltinger, U. & Wachsmuth, I. (2011). Harvesting Wikipedia knowledge to identify topics in ongoing natural language dialogs. In *Proceedings of the 2011 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2011)* (S. 445-450). Lyon, France.
- Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.
- Bublitz, W. (1989). Topical coherence in spoken discourse. *Studia Anglica Posnaniensia*, 22, 31-51.
- Button, G. & Casey, N. (1985). Topic nomination and topic pursuit. *Human Studies*, 8 (1), 3-55.
- Cassell, J., Bickmore, T., Billinghamurst, M., Campbell, L., Chang, K., Vilhjálmsón, H. et al. (1999). Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'99)* (S. 520-527). Pittsburgh, Pennsylvania, USA.
- Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsón, H. & Yan, H. (2000). Human conversation as a system framework: Designing embodied conversational agents. In J. Cassell, J. Sullivan & E. Churchill (Hrsg.), *Embodied Conversational Agents* (S. 29-63). MIT Press.
- Chafe, W. L. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.
- Chahine, C. A., Chaignaud, N., Kotowicz, J.-P. & Pécuchet, J.-P. (2011). Conceptual indexing of documents using Wikipedia. In *Proceedings of the 2011 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2011)* (S. 195-202). Lyon, France.

- Chen, H. & Dumais, S. (2000). Bringing order to the web: Automatically categorizing search results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'00)* (S. 145-152). The Hague, The Netherlands.
- Chernov, S., Iofciu, T., Nejdl, W. & Zhou, X. (2006). Extracting semantic relationships between Wikipedia categories. In *1st International Workshop: From Wiki to Semantics (SemWiki 2006), co-located with ESWC 2006*. Budva, Montenegro.
- Clark, H. H. (1996). *Using Language*. Cambridge University Press.
- Clark, H. H. & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13, 259-294.
- Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82 (6), 407-428.
- De Carolis, B., Rosis, F. de, Carofiglio, V., Pelachaud, C. & Poggi, I. (2001). Interactive information presentation by an embodied animated agent. In *International Workshop on Information Presentation and Natural Multimodal Dialogue*. Verona, Italy.
- Dengel, A. (2012). *Semantische Technologien: Grundlagen - Konzepte - Anwendungen*. Springer.
- Dijk, T. A. van. (1977). *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. Longman.
- Downing, A. (2000). Talking topically. *CIRCLE of Linguistics Applied to Communication (CLAC)*, 3, 31-50.
- Duffy, S. A., Morris, R. K. & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27 (4), 429-446.
- Ellis, J. (1966). On contextual meaning. In *In Memory of J. R. Firth*. Longmans.
- Endrass, B., Rehm, M. & André, E. (2011). Planning small talk behavior with cultural influences for multiagent systems. *Computer Speech & Language*, 25, 158-174.
- Ervin-Tripp, S. (1972). On sociolinguistic rules: Alternation and co-occurrence. In J. J. Gumperz & D. Hymes (Hrsg.), *Directions in Sociolinguistics: The Ethnography of Communication*. Holt, Rinehart and Winston.
- Feldman, R. & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.

- Fellbaum, C. (2010). WordNet. In R. Poli, M. Healy & A. Kameas (Hrsg.), *Theory and Applications of Ontology: Computer Applications* (S. 231-243). Springer.
- Foster, M. E. (2007). Universal access in human-computer interaction. Ambient interaction. *Lecture Notes in Computer Science*, 4555, 828-837.
- Foster, M. E. & Matheson, C. (2008). Following assembly plans in cooperative, task-based human-robot dialogue. In *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue (Londial 2008)*. London, England.
- Gabrilovich, E. & Markovitch, S. (2007). Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)* (S. 1606-1611). Hyderabad, India.
- Gängler, B. & Wachsmuth, I. (1993). Sprachverstehende Computer - Maschinelle Verarbeitung der deutschen Sprache in der Künstlichen Intelligenz. *Forschung an der Universität Bielefeld*, 7, 14-18.
- Gardner, R. (1987). The identification and role of topic in spoken interaction. *Semiotica*, 65 (1-2), 129-141.
- Geluykens, R. (1999). It takes two to cohere: The collaborative dimension of topical coherence in conversation. In W. Bublitz, U. Lenk & E. Ventola (Hrsg.), *Coherence in Spoken and Written Discourse* (S. 35-53). John Benjamins Publishing.
- Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, 438, 900-901.
- Glina, E. M. & Kang, B.-H. (2010). Conversation system with state information. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 14 (6), 741-745.
- Goodenough, D. R. & Weiner, S. L. (1978). The role of conversational passing moves in the management of topical transitions. *Discourse Processes*, 1, 395-404.
- Goutsos, D. (1997). *Modeling Discourse Topic: Sequential Relations and Strategies in Expository Text*. Ablex Publishing Corporation.
- Haasch, A., Hohenner, S., Hüwel, S., Kleinhagenbrock, M., Lang, S., Toptsis, I. et al. (2004). BIRON – The Bielefeld robot companion. In *Proceedings of the International Workshop on Advances in Service Robotics* (S. 27-32). Stuttgart, Germany.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. Longman.
- Hamp, B. & Feldweg, H. (1997). GermaNet - a lexical-semantic net for German. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*. Madrid, Spain.

- Helbig, H. (2006). *Knowledge Representation and the Semantic of Natural Language*. Springer.
- Helfman, J. I. & Isbell, C. L. (1995). *Ishmail: Immediate identification of important information* (Bericht). AT&T Labs.
- Heritage, J. C. & Watson, D. R. (1979). Formulations as conversational objects. In G. Psathas (Hrsg.), *Everyday Language: Studies in Ethnomethodology* (S. 123-162). Irvington.
- Herzog, O. & Rollinger, C.-R. (Hrsg.). (1991). *Text Understanding in LILOG: Integrating Computational Linguistics and Artificial Intelligence*. Springer (LNAI 546).
- Hesse, W. (2002). Ontologie(n). *Informatik Spektrum*, 25 (6), 477-480.
- Heyer, G., Quasthoff, U. & Wittig, T. (2006). *Text Mining: Wissensrohstoff Text - Konzepte, Algorithmen, Ergebnisse*. W3L-Verlag.
- Hobbs, J. (1990). Topic drift. In B. Dorval (Hrsg.), *Conversational Organization and Its Development* (S. 3-22). Ablex Publishing.
- Howe, M. (1991). Collaboration on topic change in conversation. *Kansas Working Papers in Linguistics*, 16, 1-14.
- Huber, M. J. (o. J.). JAM agents in a nutshell.
- Hurtig, R. (1977). Toward a functional theory of discourse. In R. O. Freedle (Hrsg.), *Discourse Production and Comprehension* (Bd. 1, S. 89-106). Ablex.
- Jurafsky, D. & Martin, J. H. (2009). *Speech and Language Processing*. Pearson Prentice Hall.
- Keenan, E. O. & Schieffelin, B. B. (1976). Topic as a discourse notion: A study of topic in the conversations of children and adults. In C. N. Li (Hrsg.), *Subject and Topic* (S. 335-384). Academic Press.
- Kellermann, K. & Palomares, N. A. (2004). Topical profiling: Emergent, co-occurring, and relationally defining topics in talk. *Journal of Language and Social Psychology*, 23 (3), 308-337.
- Kienreich, W. & Strohmaier, M. (2006). Wissensmodellierung – Basis für die Anwendung semantischer Technologien. In T. Pellegrini & A. Blumauer (Hrsg.), *Semantic Web – Wege zur vernetzten Wissensgesellschaft* (S. 359-371). Springer.
- Kintsch, W. & Dijk, T. A. van. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85 (5), 363-394.

- Klyne, G. & Carroll, J. (2004). *Resource description framework (RDF): Concepts and abstract syntax*. <http://www.w3.org/TR/rdf-concepts/>. (Zuletzt aufgerufen am 22. Oktober 2012)
- Kohlenberg, K. (2006). Die anarchische Wiki-Welt. In *DIE ZEIT* (Bd. vom 07.09.2006).
- Kopp, S. (2003). *Synthese und Koordination von Sprache und Gestik für virtuelle multimodale Agenten*. Akademische Verlagsgesellschaft.
- Kopp, S., Gesellensetter, L., Krämer, N. & Wachsmuth, I. (2005). A conversational agent as museum guide – Design and evaluation of a real-world application. In *Proceedings of Intelligent Virtual Agents (IVA 2005)* (S. 329–343). Springer.
- Krüger-Thielmann, K. & Paijmans, H. (2004). Informationserschließung. In H. Lobin & L. Lemnitzer (Hrsg.), *Texttechnologie: Perspektiven und Anwendungen*. Stauffenburg.
- Landis, M. H. & Burt, H. E. (1924). A study of conversations. *The Journal of Comparative Psychology*, 4, 81-89.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press.
- Linell, P. (1998). *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical Perspectives*. John Benjamins Publishing Co.
- Linke, A., Nussbaumer, M. & Portmann, P. R. (2004). *Studienbuch Linguistik* (A. Burckhardt, A. Linke & S. Wichter, Hrsg.). Niemeyer.
- Litman, D. J. & Allen, J. F. (1987). A plan recognition model for subdialogues in conversations. *Cognitive Science*, 11, 163-200.
- Maas, J. F. (2007). *Dynamische Themenerkennung in situierter Mensch-Roboter-Kommunikation*. Universität Bielefeld.
- Manning, C. D., Raghavan, P. & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Manning, C. D. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.
- Mattar, N. & Wachsmuth, I. (2012). Who are you? On the acquisition of information about people for an agent that remembers. In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART 2012)* (S. 98-105). Vilamoura, Portugal.

- Maynard, D. W. & Zimmermann, D. H. (1984). Topical talk, ritual and the social organization of relationship. *Social Psychology Quarterly*, 47 (4), 301-316.
- McCandless, M., Hatcher, E. & Gospodnetić, O. (2010). *Lucene in Action* (2. Aufl.). Manning.
- McCarthy, J. (1959). Programs with common sense. In *Proceedings of the Teddington Conference on the Mechanization of Thought Processes* (S. 75-91).
- McTear, M. F. (2004). *Spoken Dialogue Technology: Towards the Conversational User Interface*. Springer London.
- MDKMedia. (2012). *Malvorlagen - Schulbilder.org*. Verfügbar unter <http://www.schulbilder.org/>
- Mehta, M. & Corradini, A. (2008). Handling out of domain topics by a conversational character. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Art (DIMEA08)* (S. 273-280).
- Meyer-Fujara, J., Puppe, F. & Wachsmuth, I. (1995). Expertensysteme und Wissensmodellierung. In G. Görz (Hrsg.), *Einführung in die Künstliche Intelligenz* (2. Aufl., Kap. 7). Addison-Wesley.
- Miller, G. A. (1985). WordNet: A dictionary browser. In *Proceedings of the First Conference of the UW Centre for the New Oxford Dictionary* (S. 25-28).
- Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, 3 (4), 235-244.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston & B. K. P. Horn (Hrsg.), *The Psychology of Computer Vision* (S. 211-277). McGraw-Hill.
- Nass, C., Steuer, J. & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '94)* (S. 72-78). Boston, Massachusetts USA: ACM Press.
- National Information Standards Organization. (2005). *Guidelines for the construction, format, and management of monolingual controlled vocabularies* [ANSI/NISO Z39.19-2005]. Published by NISO Press.
- Nguyen, N. & Wachsmuth, I. (2011). From body space to interaction space - modeling spatial cooperation for virtual humans. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)* (S. 1047-1054).

- Nishida, T. (2007). Introduction. In T. Nishida (Hrsg.), *Conversational Informatics: An Engineering Approach* (Kap. 1). Wiley.
- Novielli, N. & Strapparava, C. (2011). Dialogue act classification exploiting lexical semantics. In D. Perez-Marin & I. Pascual-Nieto (Hrsg.), *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices* (S. 80-106). IGI Global.
- Oviatt, S. & Adams, B. (2000). Designing and evaluating conversational interfaces with animated characters. In J. Cassell, J. Sullivan & E. Churchill (Hrsg.), *Embodied Conversational Agents* (S. 319-345). MIT Press.
- Patel, R., Leuski, A. & Traum, D. (2006). Dealing with out of domain questions in virtual characters. In *Proceedings of the 6th international conference on Intelligent Virtual Agents (IVA'06)* (S. 121-131).
- Poetzsch, E. (2006). *Information-Retrieval: Einführung in Grundlagen und Methoden*. Poetzsch.
- Poggi, I. & Pelachaud, C. (2000). Performative facial expressions in animated faces. In J. Cassell, J. Sullivan & E. Churchill (Hrsg.), *Embodied Conversational Agents* (S. 155-188). MIT Press.
- Ponzetto, S. P. (2010). *Knowledge Acquisition from a Collaboratively Generated Encyclopedia*. IOS Press.
- Rath, R. (1979). *Kommunikationspraxis*. Vandenhoeck & Ruprecht.
- Rehm, G. (2004). Texttechnologie und das World Wide Web – Anwendungen und Perspektiven. In H. Lobin & L. Lemnitzer (Hrsg.), *Texttechnologie: Perspektiven und Anwendungen*. Stauffenburg.
- Rehm, M., Nakano, Y., Huang, H.-H., Lipi, A. A., Yamaoka, Y. & Grüneberg, F. (2008). Creating a standardized corpus of multimodal interactions for enculturating conversational interfaces. In *Proceedings of the IUI-Workshop on Enculturating Interfaces (ECI)*.
- Richard, J. C. (1980). Conversation. *Tesol Quarterly*, 4 (4), 413.
- Sacks, H., Schegloff, E. A. & Jefferson, G. (1974, Dezember). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50 (4), 696-735.
- Salton, G., Wong, A. & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18 (11), 613-620.

- Schank, R. C. (1977). Rules and topics in conversation. *Cognitive Science*, 1 (4), 421-441.
- Schneider, K. P. (1988). *Small Talk: Analysing Phatic Discourse*. Hitzeroth.
- Schönhofen, P. (2006). Identifying document topics using the wikipedia category network. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)*.
- Schreuder, R. & Flores d'Arcais, G. B. (1989). Psycholinguistic issues in the lexical representation of meaning. In W. Marslen-Wilson (Hrsg.), *Lexical Representation and Process*. MIT Press.
- Schubert, L. (2006). Turing's dream and the knowledge challenge. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI 2006)* (S. 1534-1538).
- Svennevig, J. (1999). *Getting Acquainted in Conversation*. John Benjamins Publishing.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.
- Taylor, D. (1986). *The elm filter system guide: What the filter program is, what it does, and how to use it* (Bericht). Hewlett-Packard Laboratories.
- Toutanova, K. & Manning, C. D. (2000). Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora* (S. 63-70).
- Tracy, K. (1984). Staying on topic: An explication of conversational relevance. *Discourse Processes*, 7, 447-464.
- Van den Berg, S. R. (1994, October). *Procmil - mail processing package*. FTP Archive. (Available FTP: Hostname: ftp.informatik.rwth-aachen.de/pub/packages/procmil/procmil.tar.gz)
- Ventola, E. (1979). The structure of casual conversation in English. *Journal of Pragmatics*, 3 (3-4), 267-298.
- Vercoustre, A.-M., Pehcevski, J. & Thom, J. A. (2008). Using Wikipedia categories and links in entity ranking. In *Focused Access to XML Documents* (Bd. 4862/2008, S. 321-335). Springer.

- Viégas, F. B., Wattenberg, M. & Dave, K. (2004). Cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04)* (S. 575-582).
- Waltinger, U. & Breuing, A. (2012). Internet-based communication. In A. Mehler, L. Romary & D. Gibbon (Hrsg.), *Handbook of Applied Linguistics* (Bd. Technical Communication, S. 533-569). Mouton de Gruyter.
- Waltinger, U., Breuing, A. & Wachsmuth, I. (2012). Connecting question answering and conversational agents – Contextualizing German questions for interactive question answering systems. *Künstliche Intelligenz*, 26 (4), 381-390.
- Waltinger, U. & Mehler, A. (2009). Social semantics and its evaluation by means of semantic relatedness and open topic models. In *Proceedings of the 2009 IE-EE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (S. 42-49).
- Wang, B. & Jiang, M. (2003). Recognition of topics from utterances in a spoken dialog system. In *Proceedings of the International conference on Natural Language Processing and Knowledge Engineering* (S. 186-191).
- Weizenbaum, J. (1966, January). ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9 (1), 36-45.
- Widdows, D. (2004). *Geometry and Meaning*. Center for the Study of Language and Information.
- Wolf, F. & Gibson, E. (2006). *Coherence in Natural Language: Data Structures and Applications*. MIT.
- Wooldridge, M. J. (2009). *An Introduction to MultiAgent Systems*. Wiley & Sons.
- Zesch, T. & Gurevych, I. (2007). Analysis of the Wikipedia category graph for NLP applications. In *Proceedings of the TextGraphs-2 Workshop (NAACL-HLT 2007)* (S. 1-8).